A Nuclear Magnetic Resonance Investigation of the Role of Structure and Dynamics in the Function of Chymotrypsin Inhibitor 2

Matthew J. Whitley

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics.

Chapel Hill
2010

Approved by:

Advisor: Dr. Andrew Lee

Reader: Dr. Sharon Campbell

Reader: Dr. Charles Carter

Reader: Dr. Nikolay Dokholyan

Reader: Dr. Gary Pielak

Abstract

MATTHEW J. WHITLEY: A nuclear magnetic resonance investigation of the role of structure and dynamics in the function of chymotrypsin inhibitor 2
(Under the direction of Andrew L. Lee)

The concept of proteins as static folded three-dimensional objects has yielded to the more sophisticated idea that proteins are better understood as dynamic entities constantly fluctuating about an average conformation rather than being restricted to any single conformation. With the knowledge that proteins sample many different conformational substates comes the question of whether the intrinsic conformational dynamics plays a role in protein function. The work presented in this dissertation begins to explore this question through the study of two model systems. The effects of subtle hydrophobic core mutations on the structure and internal dynamics of the nonallosteric protein chymotrypsin inhibitor 2 (CI2) were characterized using advanced nuclear magnetic resonance methodologies. These experiments revealed that the dynamics of methyl-bearing side chains respond globally and uniformly to hydrophobic core mutations without apparent regard for the position or chemical nature of the mutation; this is a clear example of long range intraprotein communication, a general property that forms the basis of the more complex phenomenon of allostery. Furthermore, the global increase in side-chain flexibility upon mutation takes place without concomitant structural perturbations. Functional assays of these mutants, however, revealed only subtle changes in inhibitory ability, and thus it appears that CI2 is not evolutionarily optimized to *harness* internal dynamics to modulate function despite its ability to *sense* perturbations through altered dynamics. The second model system studied is

chemotaxis protein Y (CheY), the response regulator domain of the bacterial chemotaxis pathway. CheY is an allosteric protein, and this work reports a characterization of the dynamics involved in phosphorylation-based allosteric activation. The initial results indicate that the classic framework for understanding CheY allostery is too rigid in its definition of distinct inactive and active conformational states. Many of the key residues used to delineate the allosteric state of CheY are dynamic on multiple time scales, thus indicating that they sample numerous conformations and precluding the definition of a sole inactive or active conformation. The results of this work will be useful as a framework for analysis of larger, more complex proteins of biomedical relevance as our theoretical understanding of and experimental capabilities for studying protein dynamics continue to increase.

Dedicated with love and appreciation to my parents, Penny and Michael Whitley, who always support me in all my endeavors and urge me to be the best person I can be, both scholastically and otherwise.

Acknowledgments

I would like to thank my advisor, Prof. Andrew Lee, for giving me the opportunity to do my dissertation research in his laboratory and for sharing his knowledge of and enthusiasm for NMR spectroscopy with me. I am certain that my experiences in the lab and the skills I have learned here will be of great benefit to me as I move forward with my career.

I must also thank Dr. Karl Koshlap and Dr. Greg Young for teaching me about NMR hardware and for their patience with my many repeated questions and mistakes during my early years in graduate school. I also appreciate the fact that I could always count on them for a stimulating conversation while working at the spectrometer.

I would like to acknowledge the additional members of my dissertation committee, Drs. Sharon Campbell, Charles Carter, Nikolay Dokholyan, and Gary Pielak, for volunteering to serve on my committee and for taking the time out of their busy schedules to attend my committee meetings over the years. Their suggestions were always challenging and useful.

Finally, I am grateful to all the members of the Lee Lab, both past and present, for creating an enjoyable environment in which to study and practice the art of scientific research. It is always easier to make progress on a project when the working environment is relaxed and congenial, and I thank my colleagues for this.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

3D	Three-dimensional
AIC	Akaike's information criterion
AP	Deuterium antiphase magnetization
BIC	Bayesian information criterion
CheY	Chemotaxis protein Y
CI2	Chymotrypsin inhibitor 2
D_z	Deuterium longitudinal magnetization
D_+	Deuterium in-phase transverse magnetization
DHFR	Dihydrofolate reductase
DQ	Deuterium double quantum coherence
HSQC	Heteronuclear single quantum correlation spectrum
μs-ms	Microsecond-to-millisecond
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser enhancement
PCR	Polymerase chain reaction
ps-ns	picosecond-to-nanosecond
QO	Deuterium quadrupolar order
RDCs	Residual dipolar couplings
S^2	Lipari-Szabo order parameter characterizing amplitude of N-H bond motions
S^2_{axis}	Lipari-Szabo order parameter characterizing amplitude of methyl symmetry axis motions
τ_e	Correlation time for internal motions

τ_m	A protein's global tumbling correlation time
WT	wild-type

I.

Introduction

I.1. Background

I.1.1. Modern view of protein function

Proteins are the workhorses of biology. In living organisms, they have a wide variety of purposes, ranging from serving as structural scaffolds for organizing the components of the cellular milieu, to directly catalyzing the complex biochemical reactions necessary for supporting life, and even to coordinating and regulating the interactions among the diverse biochemical pathways. Because of their vital role in maintaining life, scientists for many decades have been keen to unlock the mechanisms by which proteins function.

The original framework for understanding protein function was in terms of protein structure. Great interest was and continues to be placed in the determination of the three dimensional (3D) structures of proteins, with the generally accepted wisdom being that structure dictates or determines function, and thus that determining 3D protein structures at atomic resolution will lead to ultimate insight into the way a protein carries out its function. However, it is now known that proteins are not static entities frozen into one conformation after translation at the ribosome. In fact, proteins actually undergo motions over time scales ranging from picoseconds and faster (bond vibrations and librations) to the μs-ms regime (allosteric conformational changes, ligand binding, and conformational changes related to chemical catalysis) to time scales of seconds and slower (protein folding and unfolding).

The existence of multi-time-scale conformational dynamics in proteins requires the establishment of a new mental model, one that replaces a static view of protein structure with a conformational ensemble. Such an ensemble is a large collection of conformational substates, all of which represent deviations of greater or lesser degree from the time-averaged structure amenable to characterization by experimental techniques such as NMR spectroscopy and X-ray diffraction. The population of each of the numerous conformational substates depends on the energy of that state compared to the energy of the average conformation. Large-scale conformational deviations are comparatively rare on account of the larger energetic barrier that must be traversed, while more subtle conformational dynamics occur continually because of the low energetic barriers separating such states.

I.1.2. Rationale for this work

With the dynamical nature of proteins across multiple time scales well established, the obvious question then becomes whether the intrinsic dynamics that create the conformational ensemble can be related directly to protein function in some way. A theoretical framework for doing so already exists (*1*). The tenets of this model are shown in Figure 1.1. A protein exists in a conformational ensemble as shown in panel A. The conformational ensemble can respond to some perturbation such as ligand binding or mutation in one of three ways. The populations of the numerous conformational substates (i) can be redistributed in such a way as to result in a new time-averaged observable conformation (panel B), (ii) redistributed such that a greater (or lesser) number of conformations are actually populated, even though the time-averaged structure does not change (panel C), or (iii) a combination of the two can occur such that both the width of the distribution and the global minimum (observed

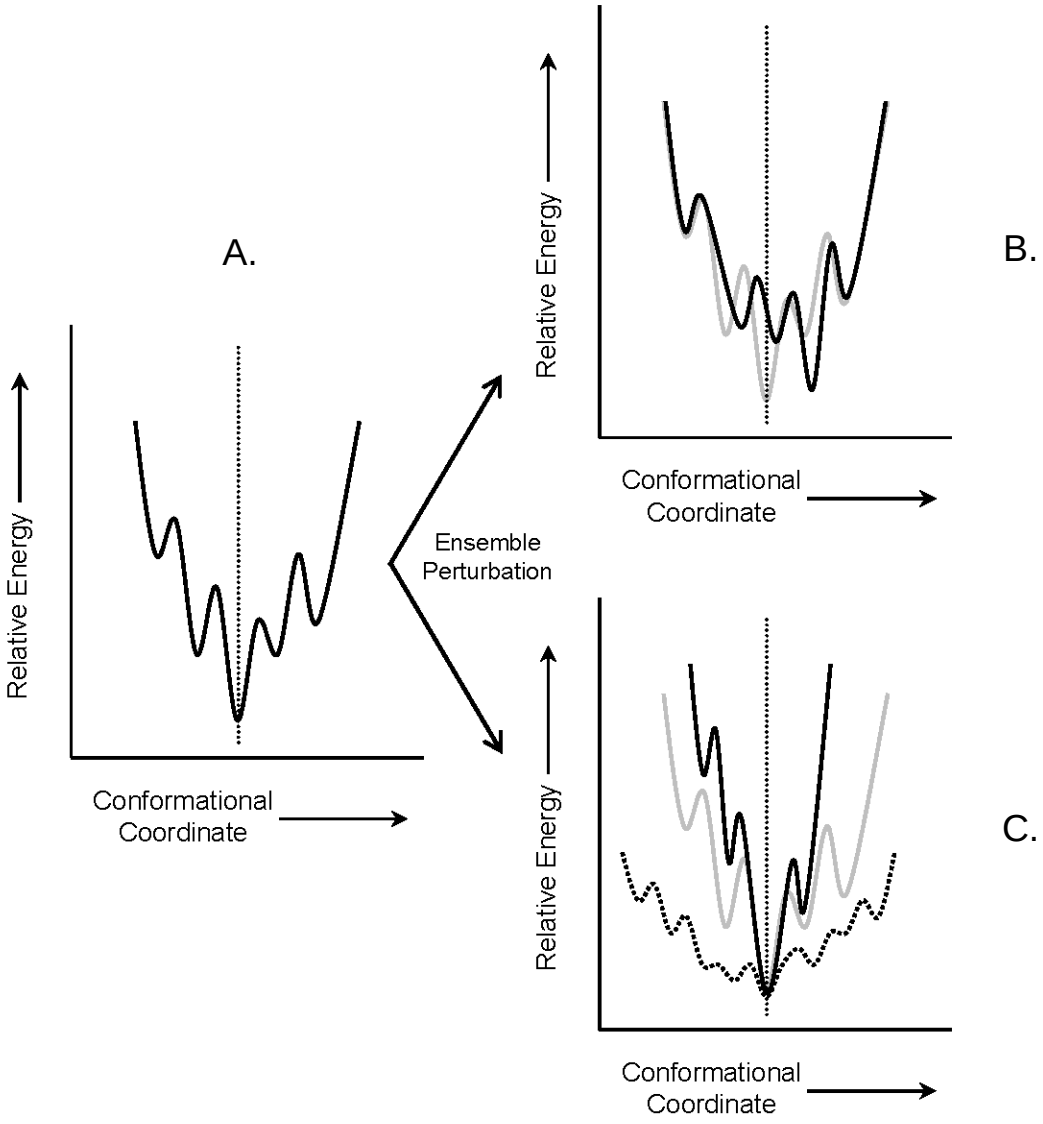


Figure 1.1. A representation of the response of the conformational ensemble to perturbation. Perturbing the system can lead to two possible outcomes. Panel A, The original distribution. Panel B, Perturbation shifts the original population distribution (gray curve) to a new distribution having a global energetic minimum at a new average conformation (black curve). Panel C, The ensemble is perturbed such that the global minimum (and thus the average conformation) of the original distribution (gray curve) does not change; rather, the width of the distribution becomes either wider (dotted curve) or narrower (black curve), representing a respective gain or loss of conformational entropy. Figure reprinted with permission from: Whitley MJ and Lee, AL. (2009) *Curr. Prot. Pept. Sci*. 10, 116-127.

structure) change. The connection to function occurs via the realization that different conformational substates can conceivably have different functional properties, with the function observed in a lab-scale experiment simply being the average of the functional ability of each of the conformational substates weighted by their respective population. Altering conformational dynamics, therefore, can theoretically impact protein function, although unequivocal demonstrations of this phenomenon are currently lacking.

One difficulty in verifying a direct connection between dynamics and function lies in the fact that still relatively few systematic characterizations of internal protein dynamics have been reported, especially in comparison to the immense body of available protein structural information. Furthermore, because the role of dynamics in function becomes most readily apparent by investigating how measurable *changes* in internal dynamics after some perturbation result in observable *changes* in protein behavior, simply characterizing the dynamics of a single protein state such as the wild type (WT) is usually insufficient. What is needed is an analysis of how the dynamics of any one protein is affected by numerous different perturbations, thus leading to an understanding of how different regions of the 3D structure modulate the dynamics of the protein as a whole.

While the number of proteins for which *any* experimental dynamics information is available remains comparatively small, the set of proteins for which a systematic experimental analysis of *changes* in dynamics in response to various perturbations is even smaller. Dynamics as a function of temperature has been studied in calmodulin (*2, 3*), ribonuclease H (*4*), adenylate kinase (*5*), and the B1 domain of protein G (*6*). Because mutations are often the source of altered protein function, there has been considerable interest in the effects of mutations on protein dynamics; among systems for which alterations in

internal dynamics have been studied as a function of mutation are calmodulin (*7*), eglin c (*8-10*), chymotrypsin inhibitor 2 (CI2) (*11*), SH3 domains (*12*), PDZ domains (*13*), and the B1 domain of protein G (*14*). Binding interactions are also a common source of modulation of protein function, and in an effort to understand how binding interactions can modulate dynamics, the effects of peptide binding on the dynamics of SH2 domains (*15*), calmodulin (*16*), Pin1 (*17*), and PDZ domains (*18*) have been investigated. A major goal in the field of protein dynamics is to clarify how dynamics might contribute to protein functions such as chemical catalysis. Such a study has been performed on dihydrofolate reductase; in this enzyme, catalytic intermediates were shown to be related dynamically to the preceding and following steps in the catalytic pathway, thus suggesting a role for dynamics in funneling DHFR through the catalytic cycle (*19*).

Even for these relatively few systems in which the effects of various perturbations on internal dynamics have been experimentally characterized, it is commonly the case that data on the corresponding effects of a given perturbation on function are not available. Thus, in order to gain insight into the relationship between dynamics and function, two complementary sets of data are needed, the first being a systematic analysis of dynamical perturbations resulting from a given type of perturbation, and the second being a measure of the functional consequences of a given perturbation. Furthermore, knowledge of the structural consequences resulting from a given perturbation is also necessary if one wishes to parse the effects on function into structural and dynamical components. Only with these three types of information can relationships among structure, dynamics, and function be examined thoroughly.

Ideally, an increased understanding of the relationship between protein dynamics and function will yield insights into the workings of biologically or pharmacologically relevant systems which can then be exploited for human benefit. Studying the dynamics-function relationship in such interesting systems, however, is fraught with difficulty, because biological functionality is often the synergistic result of multiple interacting phenomena such as allosteric regulation, conformational rearrangements both gross and fine, physical interactions with binding partners, and chemistry. The complexity of these individual phenomena often makes it difficult to parse the protein's overall functionality into contributions from individual factors such as internal dynamics.

I.1.3. Model systems studied in this work

We have therefore chosen to simplify the problem by studying two model systems of reduced complexity in order to bypass some of the pitfalls which can potentially complicate such an analysis. The main body of this work (Chapters II and III) concerns structural, dynamical, and functional studies of the model protein chymotrypsin inhibitor 2 (CI2). CI2, which comes from barley plants (*H. vulgare*), is a 7.4 kDa inhibitor of serine proteinases such as chymotrypsin and subtilisin (*20*). It is well known in the protein folding community, as its folding properties have been studied extensively via protein engineering experiments and computational techniques (*21-23*). CI2 lends itself to an analysis of the role of dynamics in function because of its simplicity – it is not an enzyme, it has no known allosteric effectors, and no conformational changes are required for function. In this dissertation, I present a systematic characterization of the effects of subtle mutations in CI2's hydrophobic core on the structure, dynamics, and function of this nonallosteric model system, the structure of which is presented in Figure 1.2A.

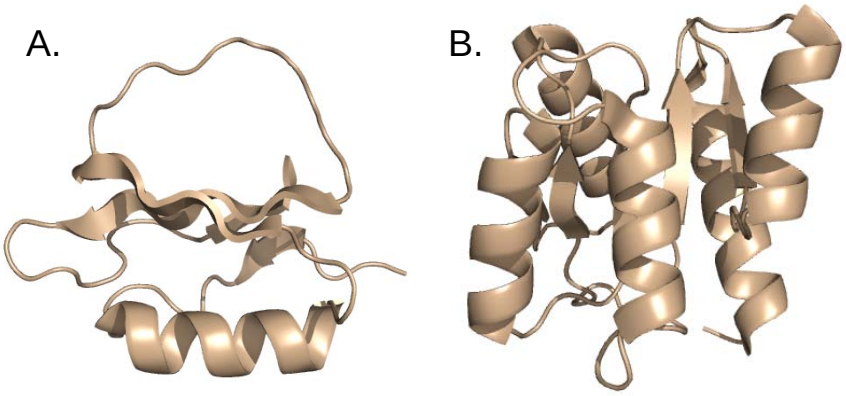


Figure 1.2. The structures of the two model proteins studied in this work. Panel A, chymotrypsin inhibitor 2 (PDB 2CI2). Panel B, chemotaxis protein Y (PDB 1FQW).

In Chapter IV, I extend the analysis to chemotaxis protein Y (CheY), a model allosteric protein. CheY is a 14.4 kDa α/β sandwich response regulator found in the *E. coli* chemotaxis pathway. It is used as a model system to study allosteric activation in proteins, because the allosteric conformational change induced by phosphorylation is relatively subtle yet results in distinct changes in chemotactic behavior; also, the fact that CheY is a single domain makes CheY a relatively simple allosteric system, since many allosteric proteins contain multiple subunits. An initial characterization of the changes in backbone and side-chain dynamics on the ps-ns and μs-ms time scales as a result of allosteric activation will be presented, as well as a structural analysis of side-chain χ_1 rotamer populations. Classical allosteric transitions are usually analyzed in purely structural terms, and much biochemical and structural information is available concerning allosteric activation in WT CheY and many mutants. However, a full understanding of allostery in this protein requires knowledge of the conformational dynamics involved in the activation process. At present, such knowledge is lacking, and Chapter IV is the first step toward closing that gap.

The remaining sections of this introductory Chapter will provide a brief introduction to the theory behind the primary nuclear magnetic resonance (NMR) methods used in this work so that the reader will be suitably equipped to understand the presented data.

I.2. NMR-based methods for characterizing protein dynamics

I.2.1. Relationship between molecular dynamics and spin relaxation

The fundamental foundation of the NMR phenomenon is the absorption of radio-frequency energy by an NMR-active nucleus, resulting in a transition from the ground state to a nuclear excited state. The excited nucleus seeks to shed the absorbed energy in order return to the favored ground state. Because of the frequencies involved in NMR spectroscopy (megahertz range), relaxation to the ground state via spontaneous emission of a photon is an extremely improbable event, and nuclear transitions therefore rely on relaxation through the stimulated, rather than spontaneous, emission of energy.

For an excited nucleus, the requirement for returning to the ground state via stimulated emission is a local magnetic field fluctuating at the frequency ν corresponding to the energy gap between the ground and excited states ($E = h\nu$). For a given NMR-active nucleus located in a protein, there are two sources of local fluctuating magnetic fields. The first source is the rotational diffusion of the protein itself in solution. As the molecule rotates, the orientation of the nuclear magnetic dipole moment changes with respect to the external field, causing a time-dependent fluctuation in the value of the total field at each nearby point in space. If the frequency of oscillation matches the resonance frequency of a nearby nucleus, that nucleus will be relaxed. The second source of local fluctuating magnetic fields inside a biomolecule is the inherent internal dynamics of the molecule itself. Internal dynamics means the random thermal displacements of a given atom from its average position within

the molecular frame of reference. These thermal fluctuations also lead to time-dependent oscillations of the magnetic field which can relax neighboring nuclei if the frequency of fluctuation is suitable. What is now needed is a way to connect relaxation caused by locally fluctuating magnetic fields to quantities measurable using NMR spectroscopy. In order to determine whether a given nucleus will be relaxed efficiently, the individual frequency components of the locally fluctuating magnetic field it experiences must be known. The bridge between relaxation theory and experimental practice is known as the spectral density function, $J(\omega)$. The spectral density function gives the density of fluctuations at a given frequency for a given bond vector and thus the power available to excite transitions at that frequency. The higher the value of $J(\omega)$ for a given angular frequency ω, the more efficiently an excited nucleus will relax to the ground state if the transition energy $E = \hbar\omega$.

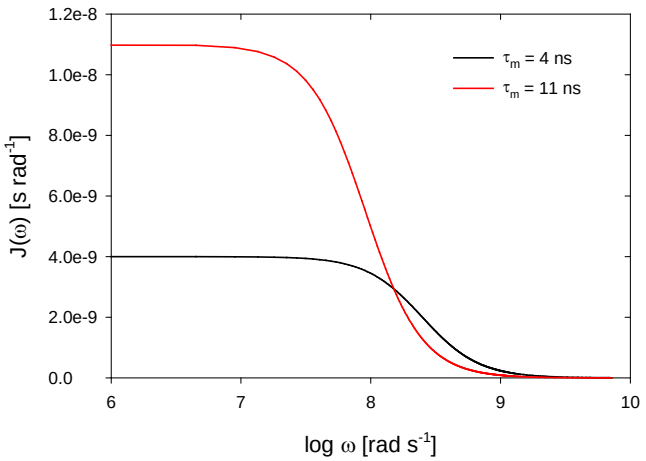


Figure 1.3. A graph of the spectral density for isotropic rotational diffusion for two proteins with different rotational correlation times. As tumbling becomes slower, the beginning of the transition of J(ω) towards zero occurs at lower ω.

Assuming that a molecule isotropically reorients in solution due to Brownian motion, the spectral density $J(\omega) = \frac{\tau_m}{1+\omega^2\tau_m^2}$, where τ_m represents the global tumbling correlation time for the protein, which is the time required for the protein to rotate 1 radian. Figure 1.3 shows the spectral density function for two hypothetical proteins with different tumbling times. The slower a protein tumbles in solution, the narrower the frequency range available to efficiently relax the excited nucleus. The above version of the spectral density function assumes that the molecule is rigid, i.e., that there is no internal

dynamics. This is clearly not the case for biomolecules, and thus the spectral density must be modified to account for both global rotational motion and internal dynamics, since both of these contribute to relaxation in real proteins in solution. The most common method for accomplishing this is known as the Lipari-Szabo model free formalism (*24*, *25*), so named because it does not make any assumptions about the physical nature of the internal dynamics other than their possible presence. The model free spectral density is given by $J_{MF}(\omega) = \frac{2}{5}\left[\frac{S^2\tau_m}{1+(\omega\tau_m)^2} + \frac{(1-S^2)\tau}{1+(\omega\tau)^2}\right]$, where $\tau^{-1} = \tau_m^{-1} + \tau_e^{-1}$. S^2 is the square of the generalized order parameter, which can vary from 0 to 1. The order parameter reflects the degree of internal motional freedom available at a given site in the protein, with a value of 0 meaning that the position can move freely and isotropically and a value of 1 meaning that the position is internally rigid and subject only to global rotational diffusion. The other new parameter, τ_e, is the motional correlation time for the internal dynamics. In effect, what the Lipari-Szabo formalism enables is the separation of dynamics processes that occur on time scales faster than overall tumbling (internal dynamics) from those that occur on time scales slower than tumbling (global rotational diffusion). Such an arbitrary partitioning of dynamic processes is not always strictly valid, and data analyzed using the model free formalism are therefore valid only under the conditions that the internal dynamics are much faster than overall tumbling ($\tau_e \ll \tau_m$) and that sub-τ_m and supra-τ_m dynamics are not correlated. These assumptions generally hold, as internal dynamics usually have τ_e values on the order of picoseconds, whereas rotational diffusion for macromolecules takes place in the nanosecond time regime. In the following sections, we turn to the application of the model free spectral density to NMR-based analyses of sub-τ_m protein backbone and side-chain dynamics.

I.2.2. Characterizing ps-ns backbone dynamics using ^{15}N spin relaxation experiments

For characterizing the dynamics of the protein main chain, the spin system comprised of the amide nitrogen and the attached proton is ideal because of its comparatively uncomplicated relaxation properties; relaxation of the amide nitrogen occurs solely through a combination of the dipolar interaction between nitrogen and its attached proton and the chemical shift anisotropy of the nitrogen itself, which results from the nonuniform spatial distribution of the electrons surrounding the nucleus. Of these two relaxation pathways, the dipolar relaxation mechanism is dominant for static magnetic field strengths at which the proton resonance frequency is less than approximately 1 GHz.

For ^{15}N nuclei, three types of data are commonly measured by NMR in order to characterize the ps-ns dynamics of the nitrogen-proton bond. These include the spin-lattice relaxation time constant T_1, the spin-spin relaxation time constant T_2, and the {^{1}H}-^{15}N steady state nuclear Overhauser enhancement (NOE). The inverses of the relaxation time constants are the corresponding relaxation rate constants. T_1 relaxation measurements report on the restoration of longitudinal (*z*-axis) magnetization to the thermal equilibrium value, while T_2 measurements report on the loss of phase coherence of magnetization located in the transverse (*x*-*y*) plane. The heteronuclear NOE measures how the populations of the ^{15}N spin are perturbed away from equilibrium when the ^{1}H spin transition is saturated (i.e., the ^{1}H populations are equalized).

Because the efficiency of relaxation depends on the value of the spectral density function, each of the above 3 relaxation parameters can be expressed in terms of weighted contributions from the spectral density sampled at various frequencies (*26*). The relevant equations relating measurable relaxation parameters to the spectral density are as follows:

$$\frac{1}{T_1} \equiv R_1 = d^2[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2 J(\omega_N)$$

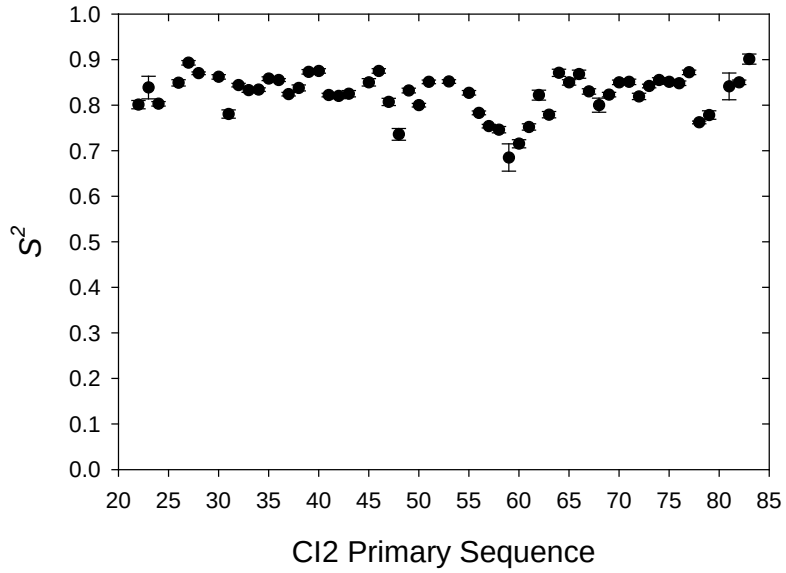
$$\frac{1}{T_2} \equiv R_2 = \frac{d^2}{2}[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)]$$

$$+ \frac{c^2}{6}[3J(\omega_N) + 4J(0)] + R_{ex}$$

$$NOE = 1 + (\gamma_H/\gamma_N)d^2[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]T_1$$

In these equations, the strength of the dipolar interaction between the ^{1}H and ^{15}N nuclei is represented by $d^2 = \frac{1}{10}\gamma_N^2\gamma_H^2\hbar^2\langle 1/r_{NH}^3\rangle^2$, and $c^2 = \frac{2}{15}\gamma_N^2 H_0^2 \Delta\sigma^2$ is a scaling factor for including the contribution of chemical shift anisotropy to the relaxation rates. R_{ex} is the contribution to R_2 from chemical exchange processes occurring in the μs-ms time regime if such processes are present. The gyromagnetic ratio for an NMR-active nucleus *w* is represented by γ_w, r_{NH}^3 represents the average internuclear distance between the amide nitrogen and the amide proton, H_0 is the strength of the applied static magnetic field, and Δσ is the anisotropy of the ^{15}N chemical shift tensor, taken to be between -160 and -170 ppm.

Figure 1.4. Backbone order parameters for N-H bonds in WT chymotrypsin inhibitor 2. The error bars, which are almost as small as the data points themselves, indicate the precision with which the measurements can be made.



Extracting Lipari-Szabo dynamics parameters from the primary relaxation data involves a per-residue computational search for the values of S^2 and τ_e in the model free spectral density that globally minimize the difference between the experimental relaxation data and theoretical

relaxation values calculated from the above equations defining the relaxation parameters in terms of spectral densities. The sample order parameter data for backbone N-H bonds in chymotrypsin inhibitor 2 shown in Figure 1.4 reflect the average trend for protein backbones, namely that they are generally rigid on the ps-ns time scale, with S^2 values distributed around 0.8. The symmetric decline in order parameter occurring from residues 53-63 indicates that these residues, which make up CI2's reactive loop (the site of binding to target serine proteases), are more flexible than the rest of the protein.

I.2.3. Characterizing ps-ns side-chain dynamics using ^2H spin relaxation experiments

Deuterium is a spin-1 particle, and thus the density matrix describing the quantum state of the particle can be described in terms of five independent operators relaxing with different rates, as described by the following equations (*27*):

$$R^Q(D_Z) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2[J(\omega_D) + 4J(2\omega_D)] \qquad \text{Longitudinal Magnetization}$$

$$R^Q(D_+) = \frac{1}{80}\left(\frac{e^2qQ}{\hbar}\right)^2[9J(0) + 15J(\omega_D) + 6J(2\omega_D)] \qquad \text{Transverse Magnetization}$$

$$R^Q(D_+D_Z + D_ZD_+) = \frac{1}{80}\left(\frac{e^2qQ}{\hbar}\right)^2[9J(0) + 3J(\omega_D) + 6J(2\omega_D)] \qquad \text{Antiphase Magnetization}$$

$$R^Q(3D_z^2 - 2) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2[3J(\omega_D)] \qquad \text{Quadrupolar Order}$$

$$R^Q(D_+^2) = \frac{3}{40}\left(\frac{e^2qQ}{\hbar}\right)^2[J(\omega_D) + 2J(2\omega_D)] \qquad \text{Double Quantum Magnetization}$$

In the above equations, $\frac{e^2qQ}{\hbar}$ is the quadrupolar coupling constant, taken to be approximately 167 kHz (*28*), and *J(ω)* is the value of the spectral density evaluated at frequency *ω*. The quadrupolar interaction with local electric field gradients dominates deuterium relaxation, and thus the quadrupolar coupling constant is present in the above equations formulating deuterium relaxation rates in terms of spectral densities.

When deuterium is randomly incorporated into proteins during expression in bacteria, side-chain methyl groups will randomly be partitioned among four isotopomers containing either zero, one, two, or three deuterons attached to the methyl carbon. For those $-^{13}CH_2D$ isotopomers, measurement of the relaxation properties of the lone deuteron allows a characterization of the dynamics of the methyl symmetry axis in terms of amplitude, S^2_{axis}, and correlation time, $\tau_{e,axis}$ (*29*). The meaning of these parameters is analogous to the backbone dynamics parameters, and the computational fitting of the values of the parameters is achieved in a similar fashion. Because these experiments require side chains to have a terminal methyl group, only the dynamics of alanine, valine, leucine, isoleucine, threonine, and methionine side chains can be characterized. It is also possible to study side-chain dynamics via methylene carbons, for example, but methyl carbons offer several advantages. Rapid rotation around the methyl symmetry axis averages the chemical shifts of the methyl protons such that only one resonance with high signal-to-noise is observed for each methyl group in a carbon-proton correlation spectrum; for methylene carbons, the two attached protons frequently have distinct chemical shifts, resulting in two crosspeaks for a single methylene carbon. Multiple crosspeaks for the same labeled carbon make the spectrum very crowded and often result in an inability to analyze the data for certain methylene carbons on account of spectral overlap. Additionally, the accurate measurement of autorelaxation rates at methylene positions is complicated by cross-correlation between ^1H-^{13}C dipolar relaxation pathways, as well as by ^1H-^1H cross-correlated relaxation (*30, 31*). Finally, methylene relaxation experiments (*32*) are known to be quite insensitive, which leads to increased uncertainty in the extracted dynamics parameters. Therefore, using methyl-based ^2H relaxation methods in conjunction with ^{15}N-based measurements to characterize both side-

chain and backbone motions helps provide a thorough picture of internal protein dynamics on the ps-ns time scale.

I.3. NMR-based methods for characterizing aspects of protein structure

Events such as ligand binding and mutation can alter protein structure. Since a knowledge of protein structure is necessary to understand protein function, it is important to characterize how protein structure changes upon perturbation. Full structure determination methods based on NMR spectroscopy and X-ray crystallography are available, but the process is time consuming. NMR methods do exist, however, which allow one to more quickly characterize aspects of protein structure without the need for solving the complete structure. The relevant methods used in this work are explained below.

I.3.1. Residual dipolar couplings probe backbone structural changes

In the previous section on relaxation measurements, the role of the dipolar interaction between two magnetic nuclei in spin relaxation was discussed. The dipolar coupling between nuclei also contains structural information in that it reports on the angular orientation between the vector connecting dipolar coupled nuclei and the external magnetic field. For two such nuclei *i* and *j*, the dipolar coupling

$$D_{ij} = -\frac{\mu_0 \gamma_i \gamma_j h}{\left(2\pi r_{ij}\right)^3} \left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle$$

where μ_0 is the permeability of free space, r_{ij} is the distance between the two coupled nuclei, and θ is the angle between the vector connecting the coupled nuclei and the direction of the NMR spectrometer's static magnetic field (*33, 34*), as shown in Figure 1.5.

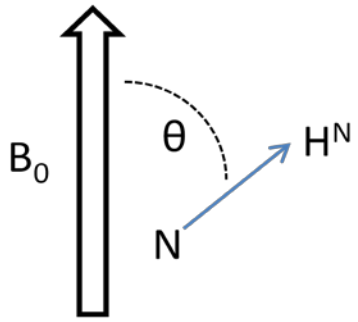


Figure 1.5. The basic geometry of residual dipolar couplings. The magnitude of the RDC varies with θ.

The brackets indicate that the angular term is a time-averaged quantity; in fact, for a macromolecule rotating isotropically in solution, the time average of this term is zero, and therefore the dipolar coupling between two nuclei does not contribute to splitting of the resonances in the NMR spectrum.

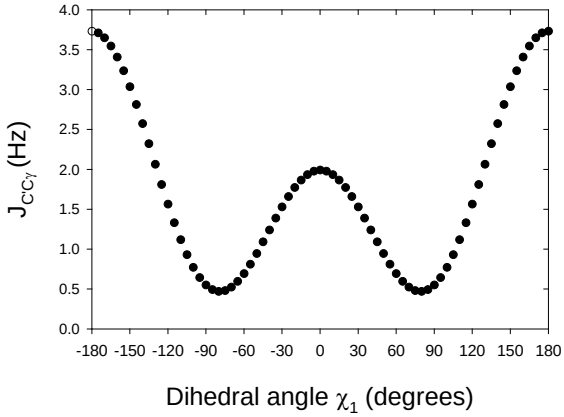
In order to measure *residual* dipolar couplings, it is necessary to introduce a small degree of alignment to the system so that there will now be a slightly preferred orientation for the molecule in solution (*35*). Creating a preferred orientation in solution is achieved by the addition of one of the various developed alignment media; these media achieve a small degree of macromolecular alignment through various means including electrostatic interactions, hydrophobic interactions, or the application of a mechanical force. In this case, the angular term will not average to zero, and NMR resonances will display a residual dipolar splitting (*33, 35*).

If a protein's structure is available, measurement of these residual dipolar splittings allows one to determine the preferred orientation of the molecule in solution. If, for example, two mutants of the same protein prefer the same orientation in solution upon addition of an alignment medium, then the RDC measured at the same position in the two mutants should be equal. However, if the structures of the two mutants are different, the RDCs of the two mutants at the same position will be different. Thus, residual dipolar couplings can be used to probe local structural differences among similar proteins aligned in the same manner. In the current work, we have measured residual dipolar couplings for backbone N-H^N and C^α-H^α vectors to study changes in backbone structure after mutating chymotrypsin inhibitor 2.

I.3.2. Use of J-couplings to study side-chain structure

While dipolar coupling is a through-space interaction, *J*-couplings (or scalar couplings) are through-bond interactions, because they arise from interactions between nuclei and the electrons in the chemical bonds connecting them. In general, scalar couplings have a measurable magnitude for nuclei separated by up to three covalent bonds. The magnitude of the scalar coupling for atoms connected by 3 chemical bonds is modulated by their dihedral angle according to the Karplus equation, $J(\phi) = A\cos^2\phi + B\cos\phi + C$ (*36*). In this equation, a dihedral angle of ϕ results in an approximate scalar coupling value of $J(\phi)$; *A*, *B*, and *C* are empirically determined constants. A graph of the Karplus curve for the dihedral angle between the protein backbone carbonyl carbon and side-chain C^{γ} is shown in Figure 1.6. By measuring the scalar coupling between these two atoms, ${}^{3}J_{C'C\gamma}$, and also the scalar coupling between the amide nitrogen and side-chain C^{γ}, ${}^{3}J_{NC\gamma}$, it is possible to calculate the populations of the three canonical χ_1 rotameric states, -60°, +60°, and 180° (*37, 38*). Knowledge of the populations of the rotameric states allows one to check for changes in side-chain structure in response to perturbations.

Figure 1.6. The Karplus curve for the dihedral angle between the backbone carbonyl carbon and side-chain γ-carbons. The values of A, B, and C in the Karplus equation are 2.31, -0.87, and 0.55, respectively.



Analogously to the RDC measurements, the rotamer populations for two mutants of a protein can be calculated, and differences in the sets of rotamer population data indicate that the mutations have had differential effects at that side-chain position. The experiments performed in this study can only be used to calculate rotamer populations for amino acids

with a γ-methyl group, meaning that rotamer populations of only valine, threonine, and isoleucine were analyzed in the present work (*39*, *40*).

II.

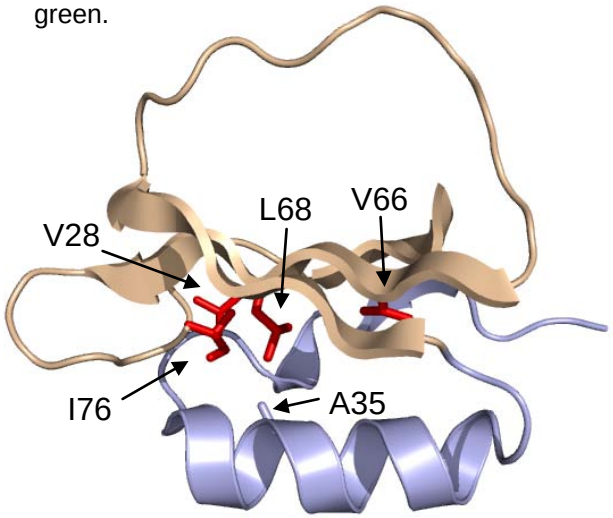
Characterization of the Structural and Dynamical Consequences of Hydrophobic Core Mutations in CI2

II.1. Introduction

The realization that internal dynamics across multiple time scales can impact both protein thermodynamics and function to a significant degree has changed the conceptual framework for phenomena such as protein-ligand interactions, catalysis, and signal transduction. Motions in the picosecond-nanosecond (ps-ns) time regime impact a protein's thermodynamics via contributions to its conformational entropy and thus total free energy (*1, 16, 41-49*), while slower motions taking place on the microsecond-to-millisecond (μs-ms) time scale have typically been linked to allostery (*50, 51*) and enzymatic catalysis (*52, 53*). Less well understood, however, are the physical determinants of protein internal motions, especially side-chain dynamics. Efforts to predict internal dynamics based on structure have shown promise (*54, 55*), but the moderate correlations observed in such attempts indicate that more than one or a few different factors must come into play.

In this study, we employ a range of sophisticated NMR methodologies to characterize the dynamical consequences of five hydrophobic core mutations in chymotrypsin inhibitor 2

Figure 2.1. The structure of chymotrypsin inhibitor 2 (PDB code 2CI2), with the sites of mutation examined in this chapter shown in red. The third canonical member of the CI2 folding nucleus, A35 (not examined in this work), is shown in gray. The backbone coloring corresponds to the two distinct regions of dynamical perturbation (N-terminal region, light blue; C-terminal region, beige) revealed by ^{15}N relaxation experiments (see text). The primary sequence of CI2 is shown below the structure, with residues 53-63 (the reactive loop) colored green.



20 MKTEWPELVGKSVEEAKKVIL
41 QDKPEAQIIVLPVGTIVTMEY
62 RIDRVRLFVDKLDNIAEVPRVG 83

(CI2), a 7.3 kDa serine protease inhibitor from barley seeds (*20, 56*). The five core mutants studied (L68V, L68A, I76V, V28A, and V66A) were selected to increase understanding of several specific issues concerning internal dynamics in hydrophobic cores (Figure 2.1). Residues L68 and I76 are known members of the CI2 folding nucleus, as determined by both experimental Φ-value analysis (*21*) and computational approaches (*22, 57*). We sought to determine whether these residues, which make important contacts that stabilize the transition state during folding, are also important for modulating CI2's internal dynamics in the native state. Furthermore, computational studies of CI2 have suggested that L68 is one of two key energetic "linchpins" connecting CI2's hydrophobic core to its reactive loop (*58*). For this reason, we have selected two L68 mutants in which the length of the side chain is progressively reduced (L→V→A) in order to determine whether the dynamical character of the L68 side chain has a role in mediating this connection. V28 is not part of the canonical CI2 folding nucleus but is located near both L68 and I76 in the hydrophobic core, and thus was chosen to reveal whether residues proximal to canonical folding nuclei show similar

dynamical behavior. V66, on the other hand, is neither part of the folding nucleus nor pointed into the same cluster of core residues formed by L68, I76, and V28. Analyzing the dynamical consequences of mutating this position adds perspective for interpreting the results of mutation to the canonical core residues. As a whole, the selected core mutations differ in location, the chemical identity of the WT residue, and the degree of severity of the mutation (although all are generally conservative), and differences between the dynamics of these mutants and WT CI2 should yield information on the relative contribution of each position to modulating the observed dynamics of the WT protein. Furthermore, the acquisition of the dynamics data for WT CI2 and hydrophobic core mutants will also allow us to compare the internal dynamics of CI2 to data available for the homolog eglin c.

II.2. Experimental Procedures

II.2.1. Protein Expression and Purification

Crystallographic evidence indicates that the first 19 amino acids of CI2 are unstructured (*20*). Consequently, experimental and computational studies of CI2 are routinely performed on a Δ1-19 truncation with residue 20 replaced by methionine, as is the case in the present work. WT CI2 DNA was generated from synthetic oligonucleotides using PCR and placed into the pET28a plasmid (Novagen). CI2 mutants were created by application of site-directed mutagenic PCR to the pET28a vector containing the WT sequence. Proteins were expressed in *E. coli* BL21 (DE3) cells (Invitrogen) grown at 37 °C to an optical density at 600 nm of 0.6-0.8 in 1 L of M9 minimal medium supplemented with 50 μg/mL kanamycin and containing 99%-$^{15}NH_4Cl$, U-^{13}C-glucose, and up to 60% $^{2}H_2O$ (v/v) as required for particular NMR samples. Expression was induced by addition of IPTG to a final concentration of 1 mM and allowed to proceed for 4-4.5 hours at 37 °C in a rotary shaker.

The cells were harvested by centrifugation at 6,000 *g* for 30 min, and the cell pellet was resuspended in 40 mL of Buffer A (25 mM Tris·HCl, pH 8) and stored at -20 °C until needed. Frozen cell suspensions were subjected to 3 rounds of freeze-thawing in a dry ice/ethanol bath, and cell lysis was completed by sonicating the sample 4 × 4 minutes. Cell debris from sonication was cleared by centrifugation at 20,000 *g* for 30 min, contaminant DNA was precipitated from the solution by the addition of polyethyleneimine to a final concentration of 0.2% (w/v), and the entire solution was allowed to stir on ice for 30 min. DNA was cleared from solution by centrifugation for 30 min at 20,000 *g*. The protein solution was then dialyzed overnight against 4 L of Buffer A. Chromatographic purification of CI2 was achieved by passing the dialyzed protein solution over an anion exchange column packed with Q-Sepharose FF resin (GE Healthcare) equilibrated in Buffer A and eluted by a linear gradient of Buffer B (2 M NaCl, 25 mM Tris·HCl, pH 8). Fractions containing CI2 (judged by the absorbance at 280 nm) were pooled, concentrated, and passed through a size-exclusion column consisting of Sephadex G-50 Superfine beads (GE Healthcare) equilibrated in NMR Buffer (20 mM KPO_4, 50 mM KCl, 0.02% NaN_3, pH 7). Pure CI2-containing fractions were again pooled, concentrated, and stored at -20 °C until needed. Typically, a 1 L culture yielded approximately 60 mg of pure protein.

II.2.2. NMR Resonance Assignments

Resonance assignment experiments were carried out on an 11.7 T Varian INOVA spectrometer equipped with a triple-resonance probe and *z*-axis pulsed field gradients using uniformly ^{13}C,^{15}N-enriched CI2 samples in NMR Buffer (10% D_2O) at pH 7. Protein concentrations were 1-2 mM. Resonance assignments for the C^{α}, C^{β}, N, and H^{N} atoms of the non-proline residues of each CI2 variant were obtained using gradient-enhanced 3D

HNCACB and CBCA(CO)NH experiments (*59*). The chemical shifts of side-chain methyl carbons (for A, V, L, I, and T) were assigned using the (H)CCH$_3$-TOCSY experiment (*60*), while methionine methyl carbons were assigned using a ^{1}H-^{13}C HMBC experiment in conjunction with the other assignment experiments. The prochirality of each side-chain methyl carbon was determined using a 10% ^{13}C-labeled sample (*61*) in conjunction with standard constant-time ^{1}H-^{13}C HSQCs. All spectra were processed using NMRPipe/NMRDraw (*62*) and analyzed using NMRView (*63*).

II.2.3. NMR Relaxation Experiments

The backbone ^{15}N T_1 and T_2 relaxation time constants and the ^{1}H-^{15}N heteronuclear NOE were measured at 25 ºC with standard experiments (*64*) on 11.7 T and 14.4 T Varian INOVA spectrometers equipped with triple resonance probes and *z*-axis pulsed field gradients. The experimental samples consisted of 2 mM uniformly ^{15}N-labeled CI2 at pH 7 in 10% D_2O NMR Buffer. For WT and L68A, the T_1, T_2, and NOE experiments were performed twice at each field strength on samples originating from independent protein expressions in order to verify the reproducibility of the experimental data. For each relaxation experiment, peak intensities as a function of relaxation time were fitted to a monoexponential decay to yield the T_1 and T_2 time constants. Error estimates for the T_1 and T_2 data were made using three duplicate relaxation delay times in each experiment, whereas the error in the calculated heteronuclear NOEs was conservatively estimated as twice the RMSD of the spectral noise. After using the backbone relaxation data to determine the global tumbling time (τ_m) of each variant, Lipari-Szabo "model free" dynamics parameters (*24*) for each residue were extracted using the program relxn2.2 (*65*) by fitting the experimental ^{15}N data to one of five common forms of the spectral density function (*66*). The appropriate form of the spectra density

function was selected on a per-residue basis by calculating both Akaike's information criterion (AIC) and the Bayesian information criterion (BIC) (*67*). For the very few cases in which the AIC and BIC recommended different forms of the spectral density, the simpler form was selected.

The ps-ns dynamics of the methyl-bearing side chains of WT and mutant CI2 were characterized at 25 ºC by monitoring the relaxation as a function of time of the following five types of ^{2}H magnetization (*27, 29*): transverse in-phase (D_{+}), longitudinal (D_{z}), transverse antiphase (AP), double quantum (DQ), and quadrupolar order (QO). ^{2}H relaxation experiments for each CI2 variant were conducted using samples of 2 mM CI2 uniformly labeled with ^{15}N and ^{13}C and 60% randomly labeled with ^{2}H in 10% D_2O NMR Buffer (pH 7) on the same spectrometers used for the ^{15}N relaxation measurements. Only signals from $-CH_2D$ isotopomers could be detected in these experiments. The transverse and longitudinal relaxation time constants were measured at both 11.7 and 14.4 T, while the AP, DQ, and QO time constants were measured only at 11.7 T. As for the ^{15}N data, the error in each ^{2}H time constant measurement was estimated using three duplicate relaxation delays. To further verify the consistency of the data, measurements of the in-phase transverse and longitudinal time constants were repeated on separately prepared samples at both field strengths for both WT CI2 and L68A. Just as for the ^{15}N data, peak intensities as a function of time for each ^{2}H relaxation experiment were fitted to a monoexponential decay to extract the associated relaxation time constant. Raw DQ time constants were then corrected according to the method of Millet *et al. (27)* to yield estimates of the pure DQ time constants. Note that this paper contains an unintended omission in the formulas given. Accounting for the influence of remote protons on the dipolar relaxation of the methyl deuteron involves scaling the

effects of remote protons on a methyl proton by the factor $8\gamma_D^2/\gamma_H^2$, where γ is the gyromagnetic ratio for a given nucleus. In the Miller paper, the additional factor of 8 is absent but has been applied properly in our calculations. The seven individual ^{2}H relaxation measurements for each residue and the overall global tumbling time τ_m derived from the ^{15}N relaxation measurements were then used to calculate model free dynamics parameters for each CI2 variant using the in-house program relxD (based on relxn2.2). RelxD generates "model free" parameters by fitting the data for each methyl group to two-parameter and ("LS-2") and three-parameter ("LS-3") forms of the model free spectral density function. The statistical F-test was then used to determine whether the LS-3 model more accurately describes the experimental data, with $\alpha = 0.05$.

II.2.4. CI2 Structural Characterization.

To elucidate the mutations' effects on CI2 backbone structure, N-H^N and C$^\alpha$-H$^\alpha$ residual dipolar couplings (RDCs) were measured using stretched 6% polyacrylamide gels (*68*) as the macromolecular alignment medium. A 2D IPAP-based experiment (*69*) was used to measure N-H^N crosspeak splittings in the absence and presence of macroscopic alignment, with the RDC value being the difference between the aligned and isotropic splitting values. C$^\alpha$-H$^\alpha$ RDCs were measured analogously using a 3D HNCA-based spectrum (*70*). RDC measurements were carried out at 25 ºC on an 11.7 T Varian INOVA spectrometer using 1-2 mM CI2 samples in 10% D_2O NMR Buffer at pH 7. To verify the quality of each variant's measured RDCs, the program REDCAT (*71*) was used to compare the experimental values to values back-calculated from a high resolution WT CI2 crystal structure (PDB 2CI2).

To probe changes in side-chain structure, NMR measurements of the three-bond scalar couplings (*39, 40*) between γ-methyl groups and their respective backbone amide nitrogens

and carbonyl carbons ($^3J_{NC}$ and $^3J_{CC'}$) were used to calculate χ_1 rotamer populations (*37*, *38*) for V, I, and T residues in the CI2 variants under study. Errors in determining the ^3J values were estimated using the RMSD of the experimental noise. The experiments were carried out at 25 °C on samples containing 2 mM uniformly ^{13}C,^{15}N-labeled CI2 in 10% D_2O NMR buffer at pH 7 using an 11.7 T Varian INOVA spectrometer equipped with a cryogenic probe and *z*-axis pulsed field gradients.

II.3. Results and Discussion

II.3.1. Mutation of folding nuclei affects backbone dynamics more universally than non-folding-nucleus mutations

In addition to WT CI2, we studied the backbone dynamics of five CI2 mutants. Mutants L68V, L68A, and I76V were chosen because they belong to the canonical CI2 folding nucleus. V28A and V66A are not members of the folding nucleus, but are located in proximity to it and were selected to provide context for interpreting the results of mutating folding nucleus residues. As a reference, the average T_1 values at 500 MHz for non-reactive-loop residues in WT, L68V, L68A, I76V, V28A, and V66A were, respectively, 0.393, 0.387, 0.389, 0.396, 0.388, and 0.395 seconds. For T_2, the averages of the 500 MHz data excluding reactive loop residues were 0.175, 0.170, 0.170, 0.165, 0.166, and 0.166 seconds. The respective 500 MHz NOE averages were 0.693, 0.698, 0.696, 0.699, 0.691, and 0.694. Typical standard errors in these measurements were approximately 1%. Based on the ^{15}N T_1, T_2, and heteronuclear NOE measurements of rigid backbone amides at two static magnetic field strengths, the globally-optimized tumbling times (τ_m) for WT CI2, L68V, L68A, I76V, V28A, and V66A, were respectively calculated to be 4.05 ns, 4.13 ns, 4.10 ns, 4.29 ns, 4.07 ns, and 4.25 ns. Because the results of the Lipari-Szabo dynamics analysis are influenced by

the type of rotational diffusion a macromolecule undergoes (*72*), we used the ^{15}N data for WT CI2 and L68A to check for the presence of rotational diffusion anisotropy. The ratio $D_{\|}/D_{\perp}$ for WT CI2 and L68A, calculated using the local D_i approach (*73*), was 1.09 and 1.12, respectively, suggesting that the use of an isotropic rotational diffusion tensor is appropriate; fitting the WT data to an axially symmetric diffusion model resulted in no significant changes in extracted dynamics parameters (data not shown). Furthermore, since the mutations studied are generally isolated from solvent and thus not expected to exert a significant influence on the mutants' hydrodynamic properties, the ^{15}N relaxation data for all CI2 variants were analyzed using the isotropic rotational diffusion model.

Fitting the ^{15}N T_1, T_2, and heteronuclear NOE measurements for WT CI2 and each mutant to the Lipari-Szabo "model free" formalism yields on a per-residue basis the parameters S^2 and τ_e, which, respectively, provide information on the amplitude and correlation time on the ps-ns time scale of the motion of a particular internuclear vector, in this case the vector between the amide nitrogen and the amide proton. The parameter S^2 ranges from 0 to 1, with a value of 1 representing a completely rigid vector and a value of 0 indicating a bond vector undergoing isotropic reorientation, while the value of τ_e is ideally greater than one order of magnitude faster than overall rotational diffusion. The S^2 data for each mutant were compared to the WT by further calculating $\Delta S^2 = S^2_{mutation} - S^2_{WT}$ for each residue; we consider significant changes to be those for which the magnitude of ΔS^2 is at least twice that of the propagated error. These data can be seen in Figure 2.2. For all mutants, one overall trend is immediately obvious, namely that most of the ΔS^2 values are negative, indicating that the backbones of the mutant proteins are slightly more flexible than the WT main chain. This effect, however, is small, as most of the changes are less than -0.05 and very few reach -

0.1.

More interesting than the magnitude of changes is the pattern of backbone dynamics changes. The two non-folding-nucleus mutants (V66A, V28A) show virtually no backbone perturbations in their N-terminal half (approximately to residue 46). In contrast, the folding nucleus mutants (L68V, L68A, I76V) show increases in backbone flexibility, many of them significant, in this same region. N-terminal changes caused by the L68 mutants cluster mainly to residues forming CI2's α-helix, with the L68V mutant perturbing the N-terminal half of the helix (residues 30-36) and the L68A mutant perturbing nearly the whole helix. The I76V mutant, in contrast, perturbs the loop immediately prior to the start of the α-helix.

The C-terminal half of CI2 (approximately after residue 46) is perturbed by all mutants and generally becomes more flexible. It is also in this region where the greatest increases in flexibility ($\Delta S^2 \approx -0.1$) are to be found. In fact, many of the residues showing changes near -0.1 are the same across the various mutants. These residues include 49, 56, 60, 67, 69, 72, and 79, and are scattered throughout the protein's structure. The larger values of ΔS^2 for the residues are in almost all cases attributable to the selection of model 5 (S^2_{fast}, S^2_{slow}, $\tau_{e,slow}$) during data fitting, which indicates that the mutations all induce more complex backbone dynamics at these residues that cannot be described by simpler forms of the spectral density that contain fewer parameters.

Collectively, the ^{15}N relaxation results indicate that CI2's backbone becomes slightly more flexible upon larger-to-smaller mutation of aliphatic side chains in the hydrophobic core. What distinguishes the various mutants from each other is their ability to perturb the N-terminal half of the protein main chain. Only the folding nucleus mutants (L68V, L68A, I76V) were able to affect the N-terminal half of the protein, with the effects localized to

CI2's α-helix and the turn leading to it. Additionally, for all mutants, certain residues in the

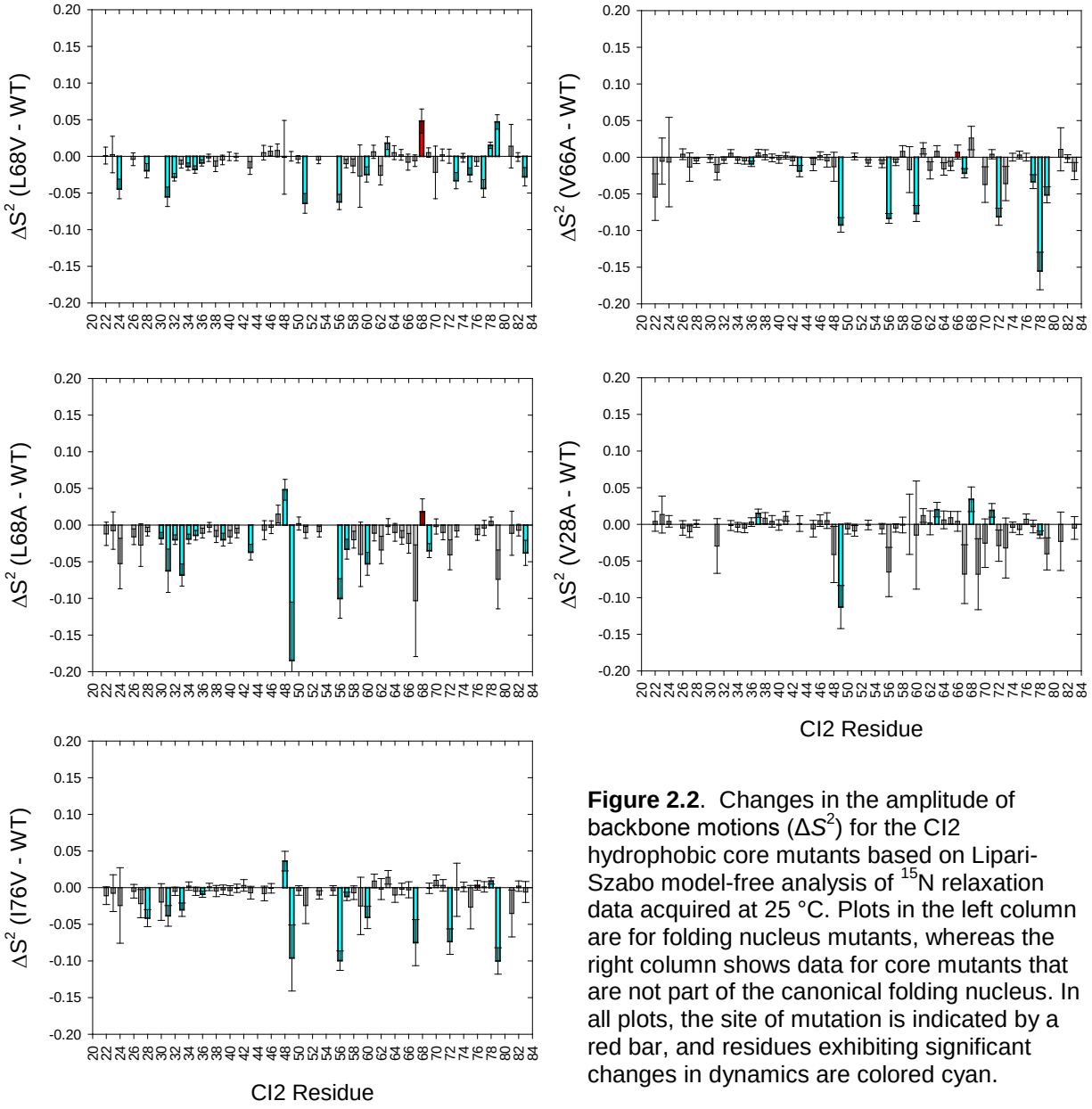


Figure 2.2. Changes in the amplitude of backbone motions (ΔS^2) for the CI2 hydrophobic core mutants based on Lipari-Szabo model-free analysis of ^{15}N relaxation data acquired at 25 °C. Plots in the left column are for folding nucleus mutants, whereas the right column shows data for core mutants that are not part of the canonical folding nucleus. In all plots, the site of mutation is indicated by a red bar, and residues exhibiting significant changes in dynamics are colored cyan.

C-terminal half appear to undergo more complex dynamics, characterized by contributions from faster and slower ps-ns motions, than in the WT. We also note that initial CPMG-based ^{15}N relaxation-dispersion experiments (*74*) did not reveal the presence of any motions taking place on the μs-ms time scale for either WT CI2 or CI2 L68A (data not shown). While all

core mutations were able to increase backbone flexibility on the ps-ns time scale, they apparently have no detectable effect on the backbone in the μs-ms time regime. Information about CI2 backbone flexibility over very slow time scales (ms - days) is available via hydrogen/deuterium exchange (HX) experiments, and HX data have been used as restraints in computational models of the CI2 native state ensemble (*75*). From these simulations, it was determined that individual RMS residue fluctuations averaged over all members of the native ensemble are mostly constant over the entire primary sequence. The profile of the largest observed fluctuations per residue, however, distinctly reveals that CI2's α-helix is the most fluctuating unit of the whole structure. Furthermore, the largest-amplitude fluctuations in the native state ensemble were observed at residue A35 and immediately before residue I76, which are two of the three folding nucleus residues. Our study indicates that only the folding nucleus mutations were able to affect ps-ns flexibility throughout the whole protein. These behaviors on different time scales are both consistent with the notion that folding nucleus residues play an important role in the native state dynamics.

II.3.2. ^{2}H-Based side-chain experiments reveal that CI2's hydrophobic core is uniquely rigid

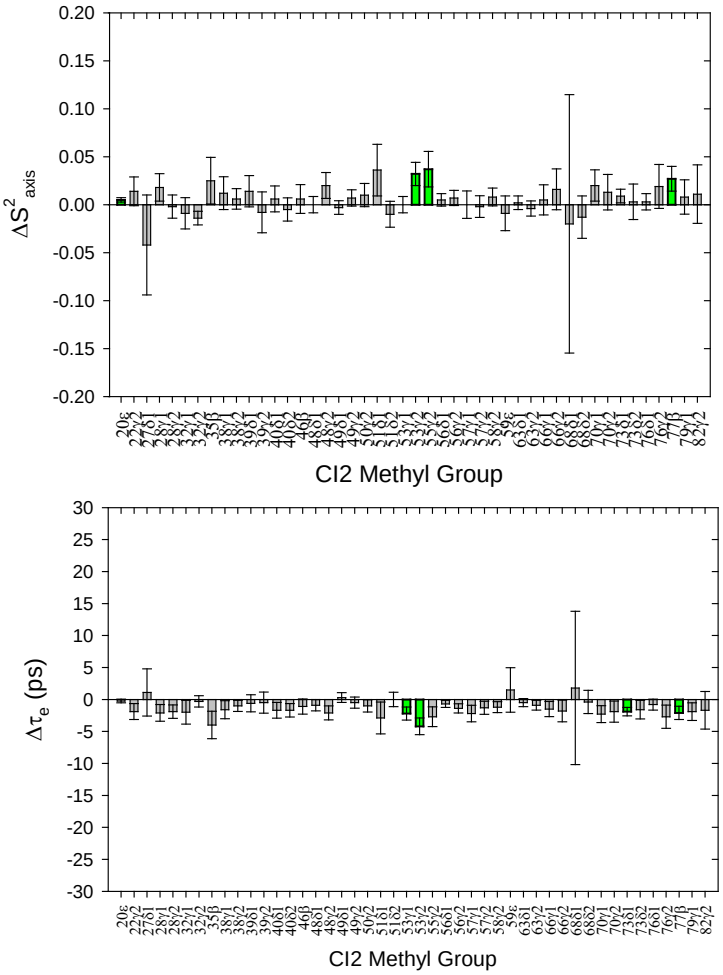
^{15}N relaxation is an excellent method for determining the flexibility of protein backbones, but the behavior of the backbone is often distinct from the side chains (*76*). It is interesting to study side-chain dynamics, because side chains are involved in key protein functions such as catalysis and molecular recognition. Studying side-chain dynamics is thus a way to gain further insight into those phenomena. ^{2}H relaxation experiments applied to $-CH_2D$ methyl isotopomers are excellent for obtaining information about the motional amplitudes and time scales of the methyl-bearing side chains in proteins. CI2 is particularly suited to analysis by these types of experiments because of the large number of methyl-containing side chains

distributed throughout its structure. Thus, CI2 contains enough $-CH_2D$ probes to reveal how core mutations affect ps-ns side-chain dynamics in all regions of the protein. Analogously to the ^{15}N relaxation studies, ^2H dynamics experiments lead to the calculation of S^2_{axis} and τ_e parameters for each methyl group. S^2_{axis} represents the amplitude of motion of the methyl group symmetry axis, and τ_e values as determined by ^2H relaxation are motional correlation times that contain contributions from rotation about and reorientation of the methyl symmetry axis, although the values are dominated by the former (*65*).

Historically, ^2H dynamics studies on methyl-bearing side chains have been conducted by measuring the relaxation of ^2H longitudinal (D_z) and transverse (D_+) magnetization as a function of time (*29*). One strength of the present study is that it expands our previous perturbation-response analyses of serine protease inhibitors (eglin c and CI2) by making use of three newer ^2H relaxation experiments in addition to the two original experiments. The major advantage of applying all five ^2H relaxation experiments (D_z, D_+, AP, QO, DQ) is that it enables robust fitting of the relaxation data to the LS-3 form of the spectral density function, which in turn allows for identification of local nanosecond-scale dynamic processes (*77*).

Seven total ^2H relaxation measurements (D_z and D_+ at two fields; AP, DQ, QO at one field) were made for WT, L68V, L68A, I76V, and V66A. For V28A, only D_+ and D_z data at two fields are available, which precludes robust fitting to the three-parameter spectral density function. One complete and independent set of data was acquired for each of two separate samples of WT CI2, and the extracted S^2_{axis} and τ_e values were compared to one another. Figure 2.3 demonstrates that the results of the two independent analyses of WT CI2 side-chain dynamics show the expected lack of change, a control that increases confidence in the

Figure 2.3. Control analysis of side-chain model-free dynamics parameters extracted from independent replicate ^{2}H relaxation measurements of different WT CI2 samples. For nearly every methyl group, the differences between the extracted S^2 axis and τ_e parameters for the two replicate sets of data are zero, within error. The very few small differences deemed significant are consistent with our 2σ standard and are colored green.



methodology before analyzing the dynamical perturbations in CI2 mutants. Changes in S^2_{axis} and τ_e between mutants and WT were calculated in Figure 2.4. Most striking is the overall pattern of changes when comparing each mutant to WT CI2. For nearly every methyl group, ΔS^2_{axis} is negative and falls in the range of -0.05 to -0.1, with an average value of approximately -0.06. This indicates that, for all mutants, the vast majority of methyl-containing side chains have become more flexible on the ps-ns time scale as a result of mutation, without regard either to the specific chemical nature of the mutation or its particular location within the hydrophobic core. This result mirrors the backbone dynamics results, but the significant increases in flexibility are more numerous and globally distributed for the side chains. Such globally-distributed, uniform changes in the amplitude of motion of nearly every detectable methyl group (Figure 2.5) were not observed either in similar studies of eglin c (*8-10*) or in mutants of the Fyn SH3 domain (*12*).

Additionally, while the amplitudes of side-chain motion are all perturbed to a similar degree seemingly irrespective of the position of mutation, the time scales of motion (reflected in the parameter τ_e) are not affected to any great degree (representative data in Figure 2.4). It thus

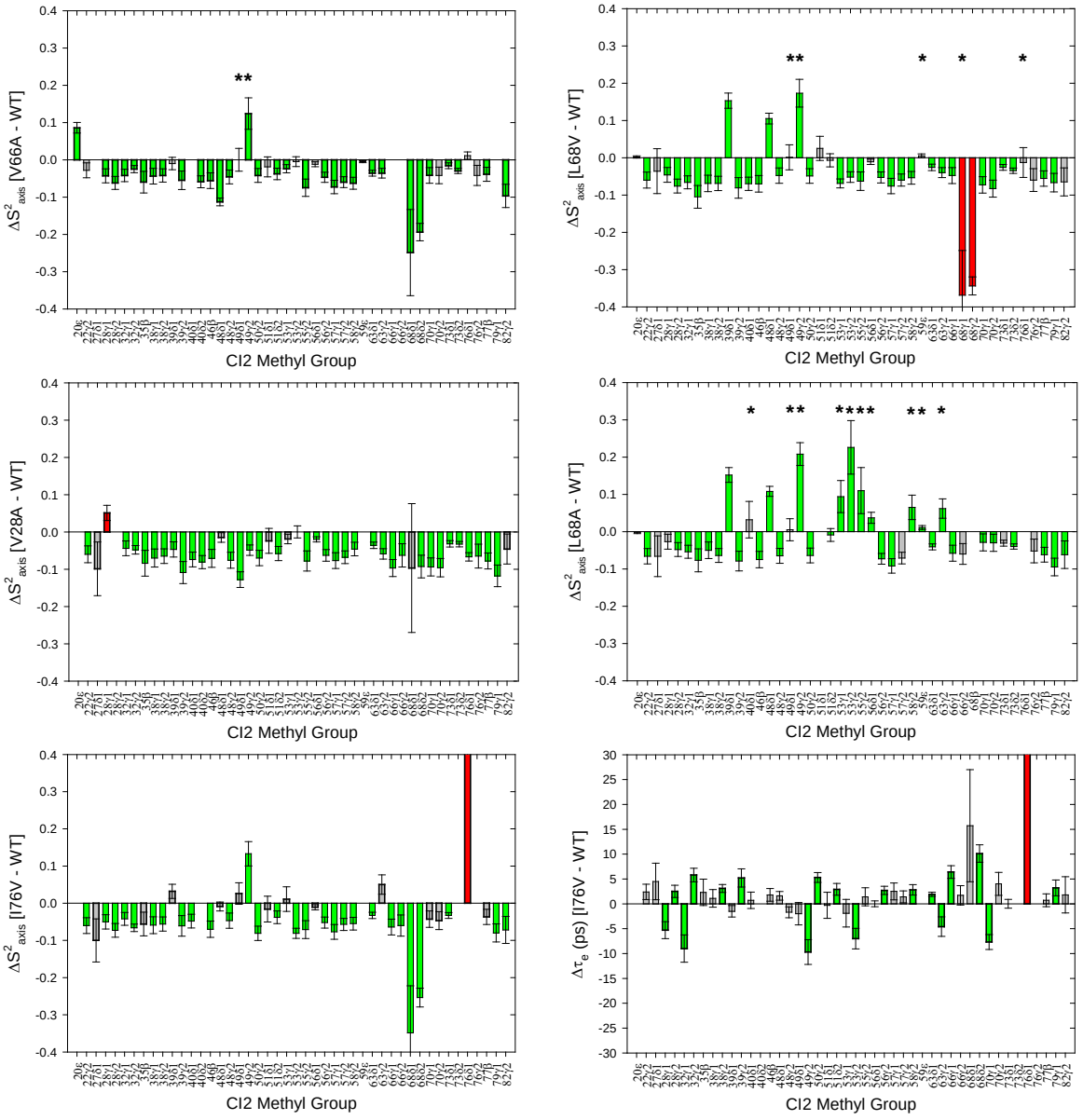
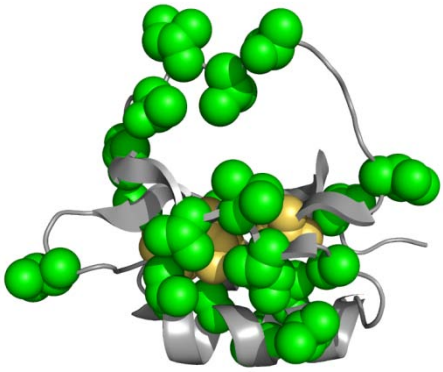


Figure 2.4. ^{2}H relaxation analysis of changes in the amplitudes of methyl-containing side-chain motions (ΔS^2_{axis}) with respect to WT CI2 for the hydrophobic core mutants studied. Green bars represent significant changes (≥ 2σ). Red bars represent the site of mutation, if data are available. The asterisks denote methyl groups experiencing slower nanosecond-scale motions based on LS-3 model selection during fitting of the relaxation data.

appears that the internal side-chain dynamics of WT CI2 are constrained at or modulated to a

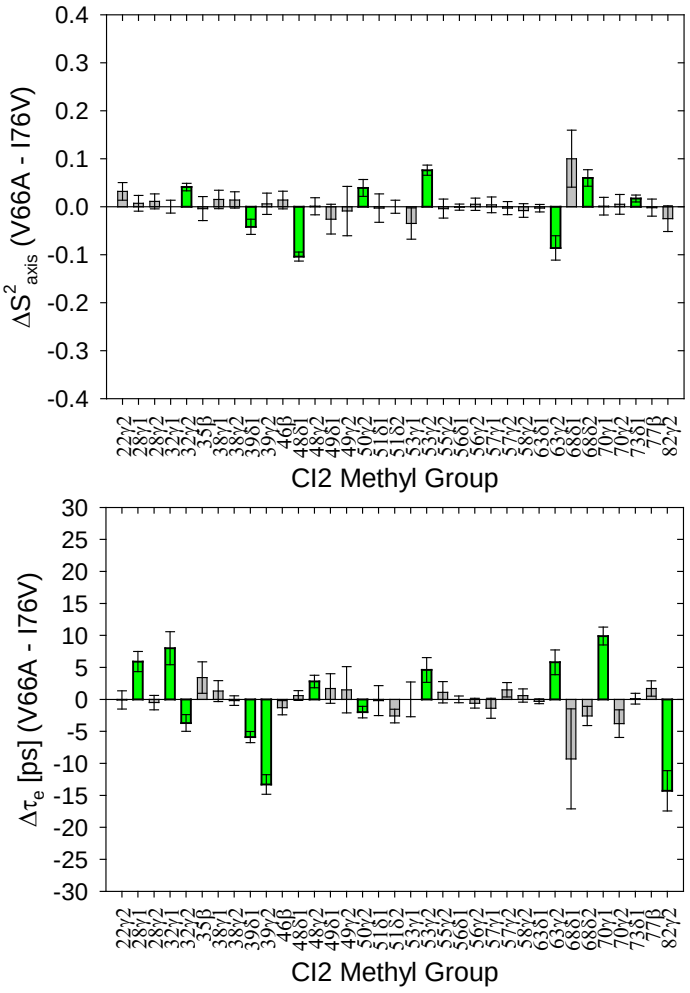
unique level of rigidity, such that mutations in the hydrophobic core result in a small but general perturbation to the physical determinants of side-chain dynamics, with the observable result being a global increase in side-chain flexibility. This global coupling of dynamics to the mutation of core residues is a novel experimental manifestation of cooperativity.

Figure 2.6. A comparison of the side-chain dynamics of a "double mutant" pair (see text) comprised of I76V CI2 and V66A CI2. The dynamics of these two mutants are much more similar to each other than either is to WT CI2.



CI2 is a classically non-allosteric protein, but if these changes in side-chain dynamics, particularly at the reactive loop, were to lead to a change in functional competency, this would by definition be an allosteric effect, and appropriate functional assays will be carried out in future studies.

Figure 2.5. A graphical summary of the ^{2}H side-chain dynamics results. In yellow are the sites of the individual single mutants (L68A, L68V, I76V, V66A, V28A). Green residues correspond to 33 methyl-bearing side chains that have significant changes in S^2_{axis} in no fewer than 4 of the 5 single mutants investigated. Changes in the amplitude of side-chain motion radiate to all regions of the three dimensional structure and encompass nearly every methyl group which could be experimentally observed.



It is important to note that we do not claim that the side chains of CI2 are abnormally rigid in an absolute sense, but rather that WT CI2's particular architecture and composition appear to support the dynamics of all the side chains such that any core perturbation "relaxes" the side-chain dynamics away from the WT parameters to a nearly identical extent, as can be seen in Figure 2.6. These are each single mutants compared to WT CI2, but are *double* mutants compared to each other. Noticeably, there are many fewer significant differences in dynamics in this double mutant comparison than there are between either of the two single mutants and WT CI2, and many of the significant differences between the two single mutants cluster to the hydrophobic core. Furthermore, among the differences that do exist there are both increases and decreases in flexibility, which is in sharp contrast to the major behavior of the single mutants compared to the WT. The lack of many substantial differences among the side-chain dynamics of the various CI2 single mutants supports our conclusion that mutation seemingly anywhere in the hydrophobic core affects the global side-chain dynamics to a very similar degree.

II.3.3. Mutation to putative energetic linchpin L68 results in increased "slower" motions at the reactive loop

The chief benefit of the expanded set of side-chain ^{2}H relaxation measurements obtained for each CI2 variant is the ability to detect the presence of "slower" nanosecond-scale motions that are not handled adequately by the standard two-parameter form of the spectral density. The data shown in Figure 2.4 resulted from model selection (two-parameter versus three-parameter spectral densities) for each methyl group in both WT CI2 and all mutants. Table 1 lists the methyl groups for which nanosecond-scale motions were detected in each variant; these methyl groups are marked with an asterisk in the panels of Figure 2.4. Except

Table 2.1. Selection of LS-3 model for CI2 mutants.

Variant	Methyl Group with Detected Nanosecond-Scale Motions
WT	$I49_{\delta}$
V66A	$I49_{\gamma2}$, $I49_{\delta}$
I76V	$I49_{\gamma2}$
L68V	$I49_{\delta1}$, $I49_{\gamma2}$, $M59_{\varepsilon}$, $V68_{\gamma1}$, $I76_{\delta1}$
L68A	$L40_{\delta1}$, $I49_{\delta1}$, $I49_{\delta2}$, $V53_{\gamma1}$, $V53_{\gamma2}$, $T55_{\gamma2}$, $I56_{\delta1}$, $T58_{\gamma2}$, $M59_{\varepsilon}$, $I63_{\gamma2}$, $V79_{\gamma2}$

for one isoleucine, the side-chain dynamics of WT CI2 are adequately described by a two-parameter model in which the fitted motional correlation times are all on a time scale much faster than overall tumbling. This is also the case for mutants V66 and I76. The series of mutations at L68 (L→V→A) reveals clearly distinct behavior. As the length of the side chain at position 68 is decreased, more and more methyl groups have a slower nanosecond-scale contribution to their dynamics. Especially for L68A, most of the residues which experience these nanosecond-scale motions in response to mutation are localized to the functional site of CI2, its reactive loop (residues 53-63). Whether these nanosecond-scale motions, which lead to apparent increased rigidity (positive ΔS^2_{axis}), will impact any functional parameters such as binding affinity for or inhibitory ability towards serine proteases will be addressed in future studies. Taken together, the model-selected side-chain dynamics data suggest that L68 is unique among core residues, as suggested by previous work (*58*). Mutations at that position alone resulted in a noticeable presence of "slower" motions, and the number of methyl groups with such motions increased as the length of the side chain at position 68 decreased. It is conceivable that the changes in reactive loop dynamics reflect the loss of the "linchpin" connection between the loop and CI2's hydrophobic core.

II.3.4. Effects of mutations on CI2 structure

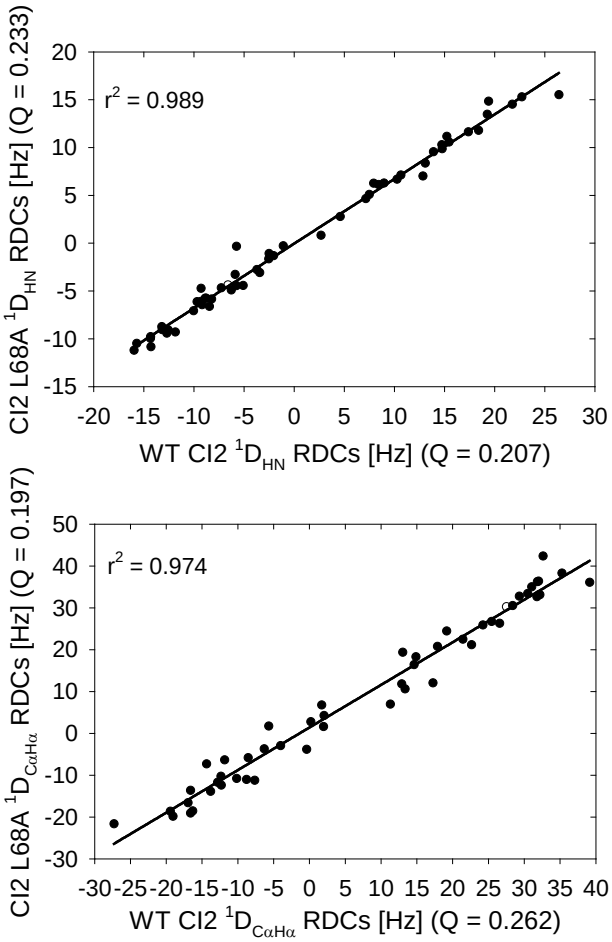


Figure 2.7. A comparison of residual dipolar couplings (RDCs) for L68A CI2 versus WT CI2. Mutant N-H (top panel) and C^{α}-H^{α} (bottom panel) RDCs correlate extremely well with WT values, indicating that no significant structural rearrangements occur upon mutation. Q-factors for all variants were calculated using the WT CI2 structure. The correlations observed here are highly representative of the mutant/WT RDC comparisons for all mutants tested.

We desired to determine the extent to which any structural perturbations caused by the mutations might be responsible for the observed widespread changes in dynamics. Structural perturbation to the backbone was assessed by measuring N-H^{N} and C^{α}-H^{α} residual dipolar couplings (RDCs). By comparing RDC data for mutants to WT data, we can assess the degree to which the backbones of the two variants are similar. Representative data are shown in Figure 2.7. The extremely high squared correlation coefficients for mutant/WT comparisons suggest that there are no gross structural perturbations caused by the mutations.

Side-chain structure was investigated for WT CI2 and the L68V and L68A mutants via measurement of three-bond scalar couplings between the backbone and side-chain γ-methyls (Val, Ile, Thr). From those measurements, we calculated χ_1 rotamer populations and compared mutant populations to those of the WT. Representative data are presented in Figure 2.8. Probing the populations of χ_1 rotameric states using scalar couplings also revealed that side-chain structure is generally conserved as well, with one notable exception. The V32 side chain (green dot in the figure)

shows an increasing preference for the +60° rotamer over the 180° rotamer as the side chain at position 68 is shortened from L to V to A. The amplitudes and time scales of the ps-ns dynamics of the V32 side chain are relatively constant between the L68V and L68A, but the rotamer populations of V32 in both mutants show signs of motional averaging, as all three rotamers are populated to some degree. Scalar couplings (and thus rotamer populations) are sensitive to motions across many time scales; since the ps-ns dynamics at V32 are relatively insensitive to mutation, but the rotamer populations show increasing changes with mutation, it is likely that some other, slower dynamic process is occurring at this residue. Examples of that type of dynamics were also found in protein L (*78*).

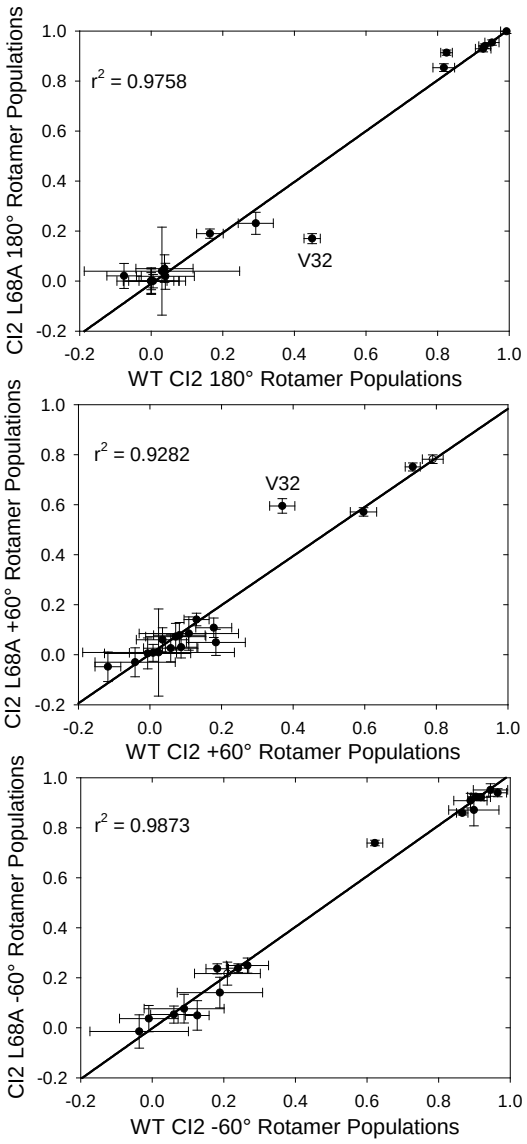


Figure 2.8. A comparison of the χ_1 rotamer populations of WT and L68A CI2 calculated from three-bond scalar coupling measurements as a probe of side-chain structure. The populations of the three canonical rotameric states (+60°, -60°, 180°) are well conserved between WT and L68A CI2, suggesting that the mutation does not grossly affect side-chain structure. The one significant exception is residue V32, which experiences a population shift from the major rotamer of 180° to the +60° rotamer as the side chain of residue 68 is progressively shortened.

In addition to ruling out gross alterations to side-chain structure, the aggregate rotamer population data suggest that the increases in mutant CI2 side-chain flexibility occur *within* the energy wells of individual

rotameric states and are not attributable to significantly increased rotamer sampling. It is known that rotameric transitions result in decreased S^2_{axis} values measured by NMR (*79, 80*), but we do not observe significant changes in rotamer populations upon mutation, except for the single example discussed above. Recently, it was shown from MD simulations that changes in dynamics within rotamer wells, rather than decreased transitioning among various rotamers, can have the greater impact on the overall change in conformational entropy upon protein-protein interaction (*42*).

II.3.5. Comparing CI2 side-chain dynamics to the homolog eglin c

An interesting open question in the field of protein dynamics is whether a given protein fold encodes the dynamics of its constituent amino acids and thus whether homologous proteins should have similar dynamics regardless of their sequence dissimilarity. By analyzing the CI2 dynamics data obtained in this study, we can contribute to the discussion of whether the dynamical character of proteins is encoded in the three-dimensional fold or whether local, sequence-specific considerations are more important. A large body of structure and dynamics information from NMR exists for the serine protease inhibitor eglin c (*9, 10*), a homolog of CI2. Similar dynamical responses upon mutating the two proteins at equivalent positions argues that the overall fold is a major determinant of internal dynamics, whereas different responses argue that sequence-specific local effects are more important. Despite the host of NMR dynamics studies published in the literature, examples of direct comparisons of the side-chain dynamics of homologous proteins are comparatively lacking. Best *et al*. compared the side-chain dynamics of homologous fnIII domains from human fibronectin and human tenascin and found the side chains in the core of the tenascin fnIII domain to be much more flexible than those in the fibronectin domain; the increased

flexibility of the tenascin fnIII side chains was attributed in part to both the amino acid composition and excess packing volume of the hydrophobic core (*81*). In contrast, other analyses of pooled side-chain dynamics data from numerous proteins suggest that, in general, there is little correlation between side-chain dynamics and structural properties such as depth of burial, solvent accessible surface area, packing density, or crystallographic B-factors (*76*).

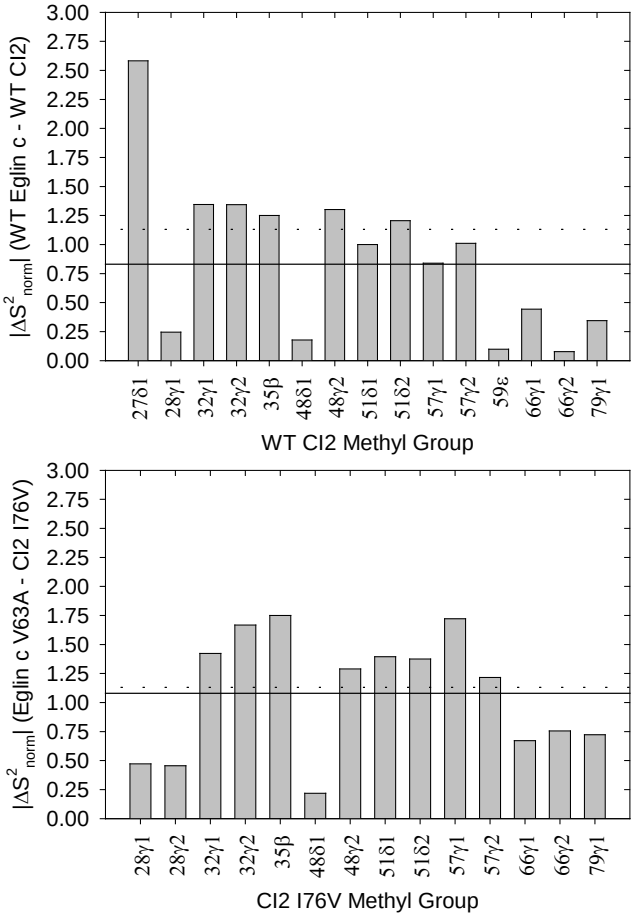


Figure 2.9. Comparisons of CI2 side-chain dynamics with respect to the homolog eglin c. Plotted are $|\Delta S^2_{norm}|$ values for WT CI2 vs. WT eglin c (top) and the homologous mutant pair I76V CI2 and V63A eglin c (bottom). The dashed line in both panels is $|\Delta S^2_{norm}| = 1.13$, the pairwise average value expected if the two proteins shared no more than random dynamical similarity. The actual pairwise averages (solid lines) are $|\Delta S^2_{norm}| = 0.83$ (top) and 1.08 (bottom).

Thus, we turned our attention to whether the side-chain dynamics of CI2 are similar to those of the homolog eglin c. Because the two proteins have different amino acids at structurally equivalent positions, the S^2_{axis} values were normalized to yield S^2_{norm} according to the method of Mittermaier *et al*. (*82*) in order to facilitate direct comparison. By calculating ΔS^2_{norm} for structurally equivalent positions in CI2 and eglin c, we can evaluate the similarity of the side-chain dynamics at those positions. In addition to WT CI2 and WT eglin c, S^2_{axis} (and thus S^2_{norm}) data at pH 7 and 25 °C exist for the mutant pair comprised of CI2 I76V and eglin c

V63A. ΔS^2_{norm} calculations for the WT/WT and mutant/mutant pairs are presented in Figure 2.9.

In order to interpret these data in terms of similarity of dynamics, it is important to appreciate what value of ΔS^2_{norm} might be expected for two proteins of no significant dynamical similarity. We approached this question by generating a Gaussian distribution of S^2_{norm} values centered at zero and having a standard deviation of one. We then selected 100,000 random pairs of S^2_{norm} values based on this Gaussian distribution, calculated the absolute value of ΔS^2_{norm} for each pair, and averaged over all the pairs to determine that two proteins of only random dynamical similarity can be expected to have a pairwise average $|\Delta S^2_{norm}|$ of approximately 1.13. This value is marked with a dashed line in Figure 2.9.

Our ability to compare CI2 and eglin c is limited by the small number of positions in the two proteins that are structurally equivalent and both occupied by amino acids with methyl-bearing side chains. While CI2 has a large number of such side chains, eglin c contains a larger degree of aromatic side chains, and thus the number of positions available for direct comparison is limited. Based on the comparisons available, however, the average $|\Delta S^2_{norm}|$ for WT CI2 versus WT eglin c is approximately 0.83, while the average value for CI2 I76V versus eglin c V63A is 1.08. This suggests semi-quantitatively that there is greater similarity between the side-chain dynamics of WT CI2 and WT eglin c than would be expected at random and would seem to support the possibility that the protein fold encodes the dynamics of its constituent amino acids. In fact, Figure 2.9 shows that most of the $|\Delta S^2_{norm}|$ values for WT CI2 and WT eglin are at worst near and at best significantly below the calculated average value for random similarity; were it not for the peculiar behavior at residue 27, the $|\Delta S^2_{norm}|$ average across all comparable methyl groups would be smaller than 0.83 and thus

indicate an even greater degree of similarity between the side-chain dynamics of the two WT proteins. The CI2 I76V and eglin c V63A mutants, however, are hardly more than randomly dynamically related to one another. This supports our proposition that native CI2 side-chain dynamics, whatever their specific physical determinants, are regulated or constrained to a unique global level and that any core mutation perturbs the side-chain dynamics away from their norm. In essence, the CI2 core mutations rob the mutants of their characteristic dynamical "CI2-ness" such that the CI2 I76V mutant shares little more than random dynamical similarity with the homologous eglin c mutant. This also indicates that the determinants of side-chain dynamics act with some subtlety, since structural characterizations revealed that there are no gross backbone or side-chain structural perturbations upon mutating CI2 core residues. These conclusions are tentative on account of the lack of available dynamics data collected under the same experimental conditions for structurally homologous positions in CI2 and eglin c.

II.4. Conclusions

This chapter presents the results of an extensive NMR study of the effects of hydrophobic core mutations in chymotrypsin inhibitor 2 on the internal dynamics of both the backbone and methyl-containing side-chains. The core mutations generally resulted in increased flexibility on the ps-ns time scale, but while the particular backbone results were mutant-dependent, the side chains showed the same slight but global increase in flexibility regardless of the position or chemical nature of the mutation. We suggest that the internal dynamics of WT CI2 are modulated in such a way as to be at some maximum of possible internal rigidity and that any small mutational perturbation in the core knocks the cohesive whole off balance, with the particular result being a small global increase in backbone and side-chain flexibility.

Additionally, a progressive shortening of the L68 side chain leads to the increasing appearance of slower nanosecond-scale dynamics, primarily localized to the reactive loop. If it is indeed the case that L68 serves as a connection between the reactive loop and the hydrophobic core, as evidence in the literature suggests, then abolishing this link by shortening the side chain might decouple the behavior of the loop from the behavior of the hydrophobic core in the ps-ns time regime.

We have also performed an initial comparison of the degree of similarity in the side-chain dynamics of two homologs, CI2 and eglin c, for both the pair of WT proteins and a pair of structurally homologous mutants. The available data suggest that the ps-ns side-chain dynamics of WT CI2 and WT eglin c exhibit more than random similarity, while the dynamics of the mutant pair are no more similar than two random proteins would be. We interpret this idea in the context of alteration to the physical determinants of side-chain dynamics upon mutation, with the CI2 I76V mutation altering these determinants in such a way that the resultant side-chain dynamics are no longer characteristic of the WT protein. Thus the similarity in the side-chain dynamics of WT CI2 and WT eglin c is abolished upon mutating two proteins at the equivalent position. The dynamical similarity of the two WT proteins suggests the importance of a protein's fold in determining internal dynamics; however, the reduced degree of similarity in the side-chain dynamics of the homologous mutants suggests that some local, sequence-dependent effect(s) from the mutations must also be contributing to the observed dynamics. Further investigation of the effects of mutation on the dynamics of homologous proteins is needed in order to clarify the relative importance of global and local influences in regulating the observed dynamics.

III.

Characterization of Structural and Functional Consequences of Mutations for CI2 in Complex with Target Serine Proteases

III.1. Introduction

The previous chapter presented a characterization of changes in both backbone and side-chain structure and dynamics on the ps-ns time scale resulting from hydrophobic core mutations in chymotrypsin inhibitor 2 and found that there are modest global changes in dynamics in the absence of significant structural changes [Whitley MJ, *et al*. Biochemistry 2008; 47:8566-8576]. In the present chapter, both the previously-studied mutants and a selection of new mutants are assayed for inhibitory ability toward the serine proteases elastase and chymotrypsin to determine the effects of mutation on function in this small model system, thus providing the functional information necessary to complete a structure-dynamics-function dataset for CI2. The assay results indicate that the mutations selected have only a subtle effect on CI2 function. The effect of complex formation on CI2 structure in solution was also analyzed, and it was found that that structural perturbations are greatest at the complex interface but also propagate toward CI2's hydrophobic core; furthermore, the structural perturbations are quite similar between WT and the L68A mutant. The combination of structural, dynamical, and functional information presented in this chapter supports the conclusion that dynamics plays at best a limited role in the function of this small model system. Based on this conclusion, the chapter ends with a discussion of possible roles for dynamics in the function of larger, more structurally and functionally complex proteins.

III.2. Materials and methods

III.2.1. Proteins

All variants of the protease inhibitors CI2 and eglin c were created, expressed, and purified as described in published work (*10*, *11*). For proteins used in NMR experiments, *E. coli* cultures were grown in media supplemented with U-^{13}C-glucose, $^{15}NH_4Cl$, and D_2O as necessary to yield suitably labeled proteins. Unlabeled proteins were used in all enzyme assays. Both bovine α-chymotrypsin and porcine pancreatic elastase were purchased from Worthington Biochemical Corporation (Lakewood, NJ, USA) in the purest grade available.

III.2.2. Determination of the apparent inhibition constant, $K_{i,app}$, for inhibition of two serine proteases

Enzyme assays were generally carried out according to a published protocol (*83*). Briefly, the steady-state rate of enzymatic hydrolysis of a chromogenic substrate was measured both before and after the addition of inhibitor by monitoring the increase in absorbance of the reaction mixture at 410 nm as a function of time. Using the slopes of the absorbance curves before and after inhibition addition as proxies for the uninhibited and inhibited reaction rates, respectively, the apparent inhibition constant was calculated from the equation $V_o/V_i = 1 + [I]/K_{i,app}$, where V_o and V_i are respective rates of substrate hydrolysis in the absence and presence of inhibitor, and [I] is the concentration of inhibitor in the assay mixture. For each inhibitor variant, multiple independent measurements of $K_{i,app}$ were made at different concentrations, and we report the results as the average value ± standard deviation.

Assays for porcine pancreatic elastase inhibition were carried out in a total volume of 1 mL of buffer comprised of 100 mM Tris (pH 8.0), 5 mM $CaCl_2$, and 0.05% v/v Tween-80. The chromogenic substrate AAA (*N*-succinyl-Ala-Ala-Ala-*p-nitroanilide*) (Sigma-Aldrich,

Inc.) was used at an initial concentration of 1 mM. In each reaction, the concentration of elastase was fixed at 1 nM, whereas the inhibitor concentration was varied through a range surrounding the approximate value of $K_{i,app}$. Assays of α-chymotrypsin inhibition were carried out similarly, except that AAPF (*N*-succinyl-Ala-Ala-Pro-Phe-*p-nitroanilide*) (Sigma-Aldrich, Inc.) was used as the chromogenic substrate, and the enzyme concentration in each reaction was 100 pM. All measurements were made using a Shimadzu UV-1601 spectrophotometer equipped with a six cell changer and a temperature controller used to maintain the reaction temperature at 25 °C. The change in absorbance was monitored for at least 45 min both before and after addition of inhibitor, and the data were analyzed using a script written for MATLAB R2008b.

III.2.3. Nuclear magnetic resonance spectroscopy

We have used NMR spectroscopy to probe the structural response of WT CI2 upon binding chymotrypsin in solution. Backbone resonance assignments of free CI2 are known from our previous work; assignments of H^N, N, C^α and C^β atoms of CI2 in complex with chymotrypsin were obtained during the present work using a combination of 3D HNCA and HNCACB spectra in conjunction with a 2D ^1H-^{15}N HSQC spectrum and the assignments of free CI2. NMR spectra were recorded on a 700 MHz Varian INOVA spectrometer equipped with a cryogenic probe and *z*-axis pulsed field gradients. NMR samples consisted of approximately 750 μM CI2 with at least a two-fold excess of protease in a buffer consisting of 75 mM HEPES, pH 8.0, 50 mM NaCl, 5 mM $CaCl_2$, and 10% D_2O v/v. The data were processed using NMRPipe/NMRDraw (*62*) and analyzed using NMRView (*63*). NMR spectra indicate that the inhibitor-protease complex was stable for approximately 15 hours before signs of degradation began to appear, consistent with known rates of CI2 cleavage by

serine proteases (*84*). The assignments for protease-bound CI2 were used along with those for free CI2 to assess and localize structural perturbations upon protease binding via a backbone chemical shift perturbation analysis. The chemical shift perturbation for each pair of backbone H^N-N backbone atoms was calculated according to the equation $\Delta\delta_{total} = \sqrt{\left(\delta_{H^N,complex} - \delta_{H^N,free}\right)^2 + \left(\frac{\delta_{N,complex} - \delta_{N,free}}{5}\right)^2}$, where $\Delta\delta_{total}$ represents the total perturbation in ppm at the backbone amide position and $\delta_{x,y}$ is the measured chemical shift for nucleus *x* in state *y*.

III.3. Results

III.3.1. Non-loop mutants show little variability in inhibitory ability as reflected in $K_{i,app}$

To explore the relationship between amino acid positions that can alter internal dynamics and function in CI2 and eglin c, we have determined the apparent inhibition constant, $K_{i,app}$, for a total of 13 CI2 variants and 5 eglin c variants with porcine pancreatic elastase and also for 5 variants of CI2 with bovine α-chymotrypsin. The location of these mutants in the 3D structure of CI2 is shown in Figure 3.1. An example of the primary cleavage velocity data used to calculate the apparent inhibition constant for each mutant can be seen in Figure 3.2. In addition to making multiple independent measurements for each variant studied, positive and negative controls were interspersed regularly with the experimental measurements. The positive control consisted of a reaction mixture containing

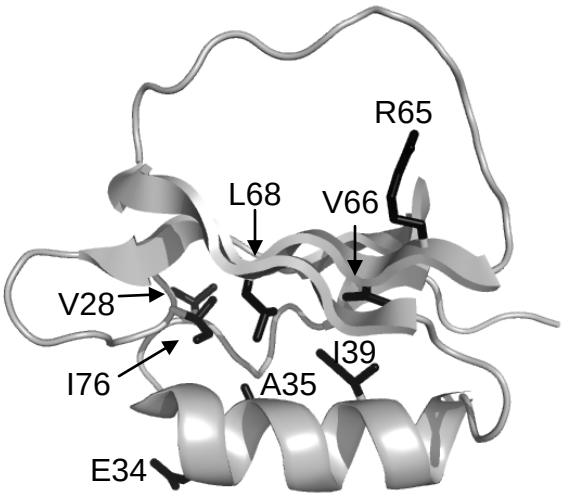


Figure 3.1. The structure of chymotrypsin inhibitor 2 (PDB 2CI2) in the free state. The positions of mutated side chains studied in the kinetics experiments are labeled and colored in black.

protease and substrate but no inhibitor; the slope of the absorbance trace remains the same throughout the experiment, assuming that there is always a large excess of substrate, and thus demonstrate that the protease does not become degraded or otherwise functionally compromised over the length of the whole measurement (Figure 3.2, reaction 6). The negative control consisted of a reaction containing substrate and inhibitor but no protease. The absorbance of such a sample should not change over time, thus ensuring that protease contamination is insignificant and that the substrate does not undergo measureable autohydrolysis over the course of the experiment.

The specific variants of each inhibitor studied were selected for particular reasons. CI2 mutants V66A, V28A, I76V, L68V, and L68A were chosen so that a possible connection between inhibitory function and changes in side-chain flexibility on the ps-ns time scale observed in our earlier work (*11*) could be probed. The other CI2 mutants were selected for their positions within the CI2 structure: E34D is located at the N-terminus of CI2's α-helix, and shortening this side chain may strain

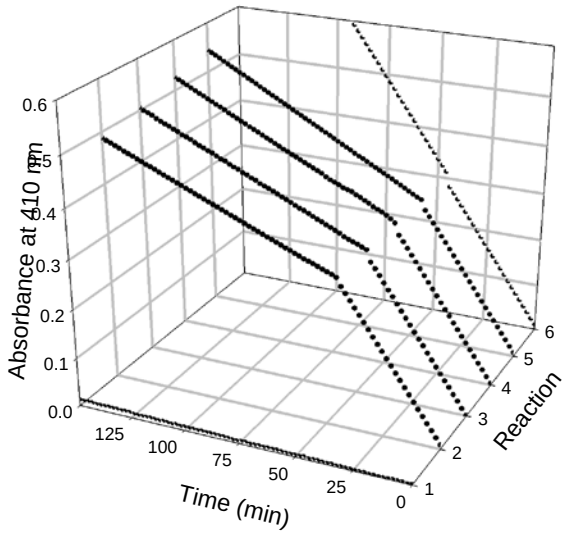


Figure 3.2. Representative data used in the determination of $K_{i,app}$ for L68A CI2 in complex with porcine pancreatic elastase. The graph shows the absorbance at 410 nm as a function of both time and inhibitor concentration. Reactions 1 and 6 are negative and positive controls, respectively. Reaction 1 contains no protease, and the absorbance therefore does not increase significantly throughout the experiment. Reaction 6 contains no inhibitor, and therefore the slope remains large and constant over the course of the measurement; in fact, the absorbance of this reaction increases to a very large value that is hidden by the scaling of the *z*-axis in this figure. L68A CI2 was added to reactions 2-5 to final concentrations of 62.5, 53, 49, and 40 nM, respectively; the different inhibitor concentrations lead to the different levels of substrate cleavage and thus to different maximum absorbances achieved during the measurement time.

electrostatic interactions at this location, A35G removes the side-chain of the most important residue of CI2's folding nucleus (*21, 22*), I39V shortens the side chain of a residue pointing into the protein's hydrophobic core, and the N-terminal truncations should weaken the formation of secondary structures at this terminus. Finally, the R65A mutation was chosen to serve as an internal control. This arginine side chain makes a critical hydrogen bond to CI2's reactive loop that is known to be critical for maintaining the loop's architecture; abolishing the hydrogen bonding capability at this position is known to severely disrupt inhibitor function in both CI2 (*85*) and *C. maxima* trypsin inhibitor V (*86*), a fellow member of the potato I inhibitor family. The mutants of eglin c under study were also chosen because of the availability of data on the dynamical consequences of the mutations (*9, 10*). Furthermore, the eglin c mutants V54A and V62A correspond to positions 68 and 76 in CI2, thus allowing the functional consequences of the mutations in the two homologous proteins to be compared and contrasted.

The results of the elastase inhibition experiments with CI2 and eglin c are presented in Table 3.1. The measured apparent inhibition constant of 12 nM for WT CI2 is similar to a previously published value of 35 nM obtained in slightly different buffer conditions (*83*). Viewed as a whole, the most striking result of the CI2 experiments is that all mutants show little variability in inhibitory ability. Most mutant $K_{i,app}$ values fall within 35% of the WT value, but the greatest measured change was a factor of 8 for the Δ3NT mutant, not considering the R65A CI2 control mutant, which is an approximately 78-fold worse inhibitor than WT CI2 on account of the structural deformations resulting from the loss of hydrogen bonding capability at this position.

A closer inspection of the CI2 data shows that the results can actually be subdivided into two groups based on the ratio of mutant to WT $K_{i,app}$. Most CI2 mutants are part of a group in which the mutant/WT ratio is at most 1.35, but the mutants L68A, Δ3NT, and R65A have

Table 3.1. Measured apparent inhibition constants, $K_{i,app}$, of variants of CI2 and eglin c in complex with porcine pancreatic elastase

Variant	$K_{i,app}$ (nM)		St. Dev. (nM)	Rel. Error (%)	Variant/WT
WT CI2	**12**	±	0.4	3.6	1.00
I76V CI2	**11**	±	2	16.0	0.88
V66A CI2	**13**	±	1	6.7	1.11
E34D CI2	**14**	±	1	9.5	1.12
V28A CI2	**14**	±	2	10.9	1.14
I39V CI2	**14**	±	1	8.6	1.16
A35G CI2	**15**	±	1	6.6	1.26
Δ2NT CI2	**16**	±	1	7.4	1.30
L68V CI2	**16**	±	0.4	2.4	1.35
L68A CI2	**24**	±	1	5.6	2.02
Δ3NT CI2	**92**	±	4	4.6	7.68
R65A CI2	**936**	±	81	8.7	77.90
WT eglin c	**10**	±	1	7.2	1.00
V54A eglin c	**7**	±	1	9.5	0.76
V18A eglin c	**15**	±	1	6.6	1.26
V34A eglin c	**13**	±	1	10.3	1.07
V62A eglin c	**14**	±	2	12.6	1.17

Notes: Δ2NT and Δ3NT represent CI2 N-terminal truncations of two and three residues, respectively. For all experiments, the listed value of $K_{i,app}$ is the average of no fewer than five independents measurements.

ratios of 2.0, 7.7, and 78.0, respectively. The increased $K_{i,app}$ for R65A is easily rationalized because of the known structural consequences of mutating this position, but the other two mutants are known to be stable, well-folded proteins (data not shown). In essence, the L68A and Δ3NT mutations, although not part of or in contact with the reactive loop, the site of binding to target serine proteases, result in a small but measurable decrease in inhibitory

ability; perturbations at these sites in the protein exert long-range effects elsewhere in the protein that alter function. CI2 is classically nonallosteric, in that it has no known allosteric effectors, but long-range intra-protein communication is the foundation of functional allostery (*87*), and it is intriguing that such subtle "allosteric-like" phenomena can be observed even in simple model systems like CI2.

Another interesting facet of the CI2 assays with elastase is the results for the two tested L68 mutations. Although the effects are small, the mutations L68V and L68A resulted in larger increases in $K_{i,app}$ than for any other single point mutation tested, excluding the control mutant. Residue L68 is thought to be a key energetic linchpin connecting CI2's hydrophobic core to its reactive loop (*58*). The mutants L68V and L68A also display dynamical properties distinct from the other mutants tested in our previous work; namely, an expanded suite of NMR-based ^{2}H relaxation measurements revealed the increasing presence of side-chain motions on the nanosecond time scale concentrated in the reactive loop as the side chain at position 68 is shortened from L to V to A (*11*); dynamical analysis of the other mutants tested (V66A, I76V, V28A) did not reveal the presence of similar levels of ns-scale side-chain dynamics. Indeed, the values $K_{i,app}$ for those three mutants with elastase are within error of the WT value, whereas the values for L68V and L68A are not, although the change at its greatest is just two-fold. Even subtle mutations at this position, as we have made in this and previous work, lead to noticeable alterations in both dynamics and function. Taken together, the structural, dynamical, and functional data available concerning position 68 all indicate that it is of special importance in this protein and appear to confirm the uniqueness of this position suggested by computational analysis (*58*). These results indicate that there is a correlation between slowing of the reactive loop dynamics, as indicated by the

increasing presence of ns-scale motions for L68 mutants, and a loss of inhibitory ability as indicated by the increase in $K_{i,app}$.

To put the elastase results into context, we assayed several mutants of CI2 against α-chymotrypsin, the results of which can be seen in Table 3.2. The numerical values of $K_{i,app}$ for WT CI2 and mutants I76V, V66A, L68V, and L68A with chymotrypsin fall in the same order as they do when assayed against elastase, with I76V yielding a very similar value to

Table 3.2. Measured apparent inhibition constants, $K_{i,app}$, of variants of CI2 in complex with bovine α-chymotrypsin

Variant	$K_{i,app}$ (nM)		St. Dev. (nM)	Rel. Error (%)	Variant/WT
WT CI2	**202**	±	17	8.4	1.00
I76V CI2	**217**	±	15	6.9	1.08
V66A CI2	**248**	±	42	17.0	1.23
L68V CI2	**259**	±	23	8.8	1.28
L68A CI2	**282**	±	9	3.1	1.40

Note: For all experiments, the listed value of $K_{i,app}$ is the average of no fewer than five independents measurements.

WT CI2, V66A being a slightly worse inhibitor, and the L68 mutants again having the greatest negative impact on inhibitory power. We note that our measured value of $K_{i,app}$ for WT CI2 in conjunction with chymotrypsin is an order of magnitude larger than the value reported in an earlier study (*83*), although the previously reported values were determined in slightly different buffer conditions and at elevated temperature (37 °C compared to 25 °C in the present study). In light of the fact that we are mostly interested in the *differences* among the assayed mutants and not the absolute values themselves, it is clear that the subtle mutations tested have only a subtle impact on CI2 inhibition of bovine α-chymotrypsin, exactly as we observed for the same experiments with pancreatic elastase.

Finally, we also measured the apparent inhibition constants of several mutants of the CI2

homolog eglin c in order to discover whether mutation at analogous positions results in similar functional consequences. All eglin c mutants were low nM inhibitors of elastase, and little difference was detected among the five variants tested, just as for CI2. Interestingly, the V54A mutant is apparently a slightly better inhibitor than WT eglin c, even though the differences are small. In eglin c, V54 occupies a position analogous to L68 in CI2. In CI2, L68 mutations resulted in less effective inhibitors, but mutating the same position in eglin c appears to have no negative impact on inhibitory function. The opposite appears to be true for the pair V62A and I76V in eglin c and CI2, respectively. I76V appears to be a neutral mutation regarding CI2 function, but V62A has the greatest impact on eglin c inhibition of elastase observed in this small set of mutants. Overall, however, it is clear that these conservative larger-to-smaller aliphatic point mutations in CI2 and eglin c have limited impact on their inhibitory ability toward porcine pancreatic elastase.

III.3.2. CI2 chemical shift perturbations upon binding chymotrypsin radiate from the reactive loop toward CI2's hydrophobic core

The history of structure-function studies on serine proteases and their small molecule and macromolecular inhibitors is long and rich, and much is known about structural determinants of binding specificity (*88*), including the importance of long-range interactions with structural elements such as loops which are not directly part of the binding site (*89, 90*). Despite the fact that proteases and protease inhibitors are classic systems in the field of protein-protein interactions, surprisingly little structural information is available for chymotrypsin inhibitor 2 in complex with chymotrypsin. Crystal structures of CI2 and mutants in complex with homologous proteases including subtilisin BPN′ (*84, 85, 91*) are available, as are structures of CI2 homologs such as

eglin c in complex with subtilisin novo (*92*) and chymotrypsin (*93*). However, there is no structure of CI2 in complex with chymotrypsin. Because of the lack of such information, we decided to apply NMR spectroscopy to perform an initial characterization of the 32 kDa CI2-chymotrypsin complex in order to gain some insight into the structural consequences of complex formation in solution. To our knowledge, this is the first NMR analysis of the complex of a macromolecular protease inhibitor in complex with a fully active WT serine protease.

Using triple resonance NMR methods, we have assigned the H^N, N, C^α, and C^β resonances of CI2 labeled uniformly with ^{13}C and ^{15}N and deuterated to approximately 80% while in complex with chymotrypsin. An overlay of the free and bound CI2 spectra is showin Figure 3.X. Using these data in conjunction with the free-state assignments obtained previously, we are in a position to assess possible structural perturbations to CI2 upon binding the much larger chymotrypsin. NMR chemical shifts are exquisitely sensitive to the local environment surrounding a particular NMR-active nucleus, and thus changes in the local environment are reflected by changed chemical shifts. A chemical shift perturbation analysis yields insight into conformational perturbations in solution without undertaking the much more involved process of solving a full solution structure; the analysis yields data on where the local environment has changed but does not reveal the exact nature or cause of the local environmental changes. Figure 3.3A shows the chemical shift perturbations experienced by WT CI2 upon binding chymotrypsin. For reference, the residues of CI2's reactive loop are shaded black in the figure, and visual inspection makes clear that the chemical shift perturbations of the largest magnitude are found at the reactive loop,

specifically with the perturbations reaching a maximum at the scissile bond (residues 59-60) and generally decreasing moving along the primary sequence in either direction. This result

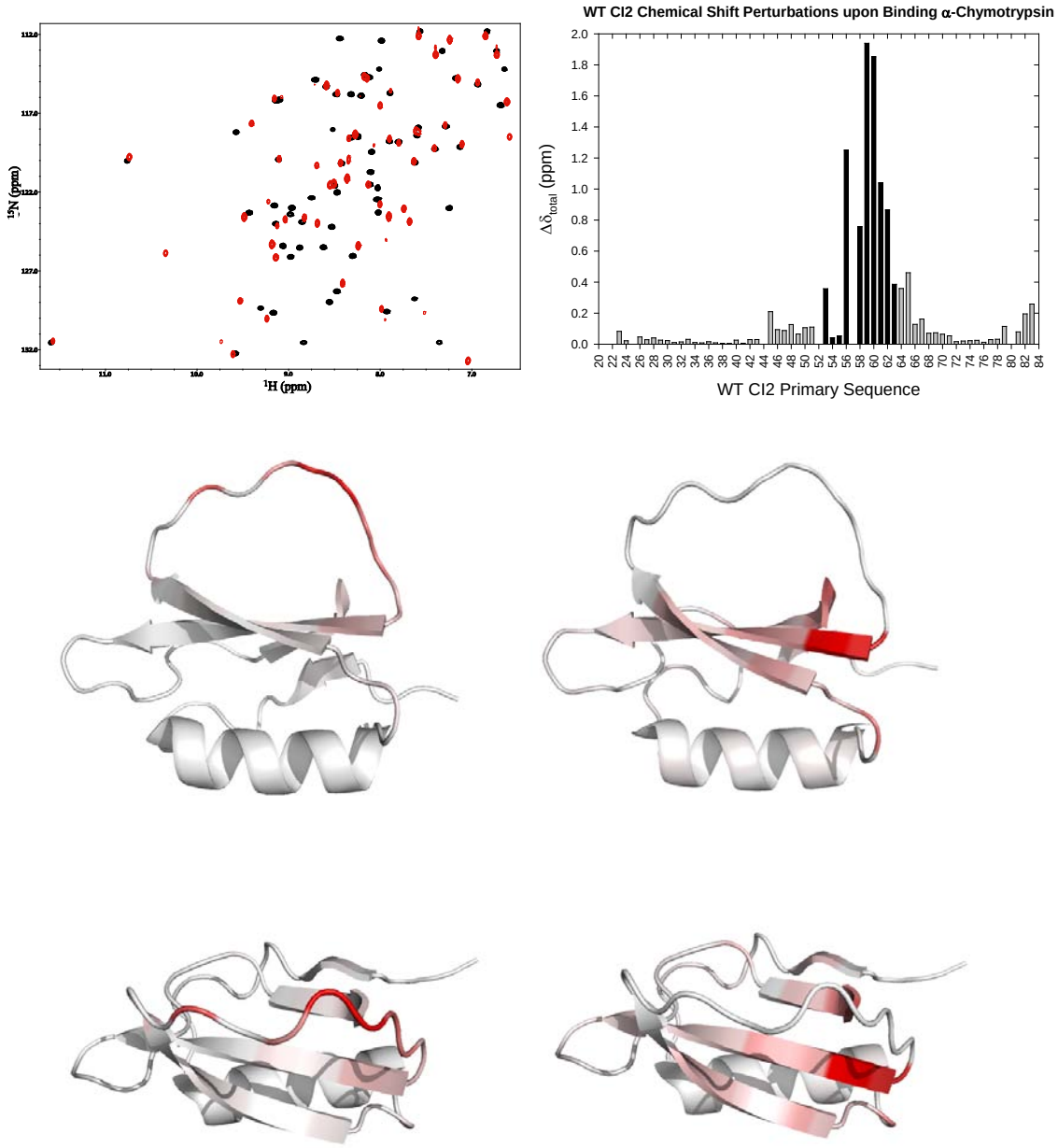


Figure 3.3. Analysis of H^N-N chemical shift perturbations in WT CI2 upon binding bovine α-chymotrypsin. Panel A shows the overlayed ^{1}H-^{15}N HSQC spectra for CI2 in the free (black) and chymotrypsin-bound (red) states. Panel B shows the calculated chemical shift perturbation as a function of CI2 residue. The black bars represent CI2's reactive loop, the site of binding to its protease targets. Panel C shows the data from panel B plotted onto the CI2 structure (PDB 2CI2). The shading from white to red makes it visually clear that the strongest perturbations are located in the reactive loop. Panel D is the same as panel C, except that the data for the reactive loop (residues 53-63) perturbations have been removed so as to reveal smaller but still significant changes that are visually masked by the dominant reactive loop perturbations in panel C.

can be rationalized by the physical nature of the system. In the free state, CI2's reactive loop protrudes from the core of the protein, thus being exposed to solution and prepared for binding to a protease target. Upon complex formation, these residues become buried, though the conformation of the reactive loop remains similar to the
contacts with chymotrypsin. Figure 3.3B shows the entire range of chemical shift free state. Thus the large perturbations in Figure 3.3 panels A and B are due primarily to new perturbations plotted onto the structure, making clear the impact of complex formation on the reactive loop of CI2. Figure 3.3C shows a re-scaled view of the chemical shift perturbations when the reactive loop residues 53-63 are excluded. This exclusion makes clearer visually that perturbations from binding at the reactive loop propagate significantly to CI2's β-sheet; in fact, binding has long distance effects, as the chemical shift perturbation at A45 is 0.2 ppm although A45 is approximately 15 Å from the closest part of the protease. Thus, binding of a protease target at the reactive loop is "sensed" in the core of the protein. The chemical shift perturbation analysis suggests that some conformational perturbation is experienced throughout the protein upon protease binding and that the force of binding propagates along the reactive loop to the core and outward to the edges of the β-sheet; of the residues in Table 3.2 that were assayed against chymotrypsin, V66 and L68 are located on the central β-strand, and these residues show the greatest change in $K_{i,app}$ for the mutants tested.

Because the L68A CI2 mutant shows the largest difference in $K_{i,app}$ of any of the single point mutations compared to the WT value for both chymotrypsin and elastase and because its side chain points from the β-strand into the hydrophobic core, we wondered whether

differential structural effects incurred upon protease binding might be responsible. To answer this question, we used the CI2 chemical shift assignments for free and bound WT and for free L68A in conjunction with a ^{1}H-^{15}N HSQC spectrum of the L68A CI2-chymotrypsin complex to assign unambiguously 45 of 64 total residues in the bound L68A mutant. The results of chemical shift perturbation calculations for the transition from free to complexed L68A CI2 are nearly identical to the chemical shift perturbations experienced by WT CI2 upon binding chymotrypsin with regard to both the overall pattern of chemical shift perturbations and the magnitude of the perturbation for each individual residue (data not shown). This suggests that differential conformational perturbations upon binding the protease are not responsible for the variations in the measured values of $K_{i,app}$ for the variants tested and thus that the dominant conformation of the free state ensemble determines inhibitory ability. The chemical shift results taken together lead us to conclude that (i) the primary structural consequences upon complex formation in solution are confined to CI2's reactive loop, but furthermore that (ii) binding the protease at the reactive loop is indeed sensed by the core of the protein, as reflected by the significant detected chemical shift perturbations throughout CI2's β-sheet. Finally, (iii) the chemical shift perturbation analysis of L68A CI2 binding to chymotrypsin resulted in nearly identical chemical shift perturbations compared to chymotrypsin complex formation with WT CI2. This result suggests that complex formation with the protease affects each of the tested CI2 variants in a similar manner.

III.4. Discussion

In this paper, we have presented the results of basic functional and structural analyses of variants of the serine protease inhibitors CI2 and eglin c in complex with the serine proteases

porcine pancreatic elastase and bovine α-chymotrypsin. As complementary information to our previously published analysis of changes in free CI2 structure and dynamics upon mutation in the hydrophobic core (*11*), we have assayed the inhibitory ability of a set of variants of the homologous serine protease inhibitors CI2 and eglin c toward the serine proteases elastase and chymotrypsin with the aim of creating a unique data set that combines site-specific dynamical, structural, and functional data to gain insight into the possible roles of dynamics and structure in the function of these small model protein systems. We have also characterized the conformational perturbations to WT and L68A CI2 upon binding chymotrypsin to determine whether differential perturbations among the various mutants upon binding a protease might also play a role in inhibitory ability.

III.4.1. Effect of dynamics on function

Overall, the functional assays revealed little significant change in the apparent inhibition constant for the tested single-site variants of CI2 and eglin c. Still, one noticeable result concerns mutations at position L68 in CI2. Our previous work showed that shortening this side chain from L→V→A led to increasing levels of ns-scale side-chain dynamics on the ps-ns time scale, a phenomenon not detected to the same degree in the other mutants studied. Interestingly, the same mutations L68V and L68A had larger effects on inhibitory ability than any other single point mutation in the hydrophobic core of CI2, even though the effects on function are a quite subtle factor of 2. However, this factor of 2 is still notable upon consideration of the fact that some classically allosteric proteins do no better; upon exposure to light, a photoallosteric PAS domain fused to DHFR resulted in an enhancement of DHFR catalysis by at best a factor of 2 (*94*). The L68 results lend credence to the previously suggested notion that L68 mediates a connection between the hydrophobic core "scaffold" of

CI2 and the functional reactive loop that binds serine protease targets. Additionally, a comparison of the results for mutants tested against both chymotrypsin and elastase reveals the same trend in apparent inhibitor constants. This suggests that these mutant inhibitors may indeed be distinct from one another with regard to functional ability, notwithstanding the only slight changes in $K_{i,app}$. Thus, the observed correlation between altered dynamics and altered function appears to be due more to ns-scale motions than to the typically faster side-chain motions detected using only the transverse and longitudinal ^{2}H relaxation experiments. Comparing the results of mutating analogous structural positions in CI2 and eglin c is also useful. When assayed for elastase inhibition, both proteins yielded similar results – apparent inhibition constants on the order of 10 nM and little change upon subtle mutation in the hydrophobic core. The value of $K_{i,app}$ for mutation at analogous positions, however, did not follow the same trend in the two homologs. This suggests that the mutations have different consequences in the two inhibitors which are significant enough, albeit subtle, to lead to distinct inhibitory abilities. The two inhibitors are known to experience distinct dynamical perturbations upon mutation at these analogous positions (*8-11*), which suggests that dynamics (or dynamics in conjunction with subtle, differential structural perturbations) does play a role in determining functional properties in these model systems. Taken as a whole, however, the fact that global changes in dynamics detected in previous work resulted in only fine changes in inhibitory ability in the variants tested leads us to believe that dynamics, at least in these conservative single-site substitutions, plays at most a subtle role in the inhibitory function of these model protein systems. Nevertheless, we did identify two CI2 mutants, L68A and Δ3NT, the effects of which were sensed at the reactive loop and resulted in at least a twofold change in $K_{i,app.}$ Long range effects are the foundation

of classical allostery, and it is therefore intriguing to find such effects present in CI2 in response to relatively subtle perturbations and despite it being a classically nonallosteric protein.

III.4.2. Chemical shift perturbations upon complex formation

Because of the exquisitely sensitive dependence of the NMR chemical shift on the local environment surrounding an NMR-active nucleus, NMR chemical shift perturbations are often used as indications of perturbations to the local environment when a protein binds another molecule. We selected WT and L68A CI2 to analyze of chemical shift perturbations upon binding chymotrypsin, as these two variants showed the largest numerical difference in $K_{i,app}$. As expected, chemical shift perturbations were large at the reactive loop, but there were also significant perturbations along the β-strands, even though they are not in direct contact with the protease, thus indicating that binding has non-local consequences for the inhibitor. Importantly, however, the chemical shift perturbation analysis for the L68A CI2 mutant in complex with chymotrypsin yielded results very similar to the WT analysis. This suggests that, whatever perturbations are experienced upon binding between WT CI2 and chymotrypsin, very similar perturbations are experienced by the L68A mutant. Thus, differential structural responses upon protease binding can be ruled out as a major determinant of functional ability. Because dynamics appears not to play a dominant role in determining function and because structural perturbations upon binding appear to be similar even among the most functionally-distinct CI2 variants, we conclude that the architecture of the CI2 reactive loop in the unbound state plays the major role in CI2 function.

III.4.3. Correspondence between structural and dynamical perturbations

We also make the intriguing observation that first third of the protein appears to be a "blind spot" to structural and dynamical perturbations. Figure 3.3A shows that region of CI2 from the N-terminus to residue 44 experiences little to no chemical shift perturbation upon protease binding, and Figure 2.2 shows that the backbone dynamics of non-folding-nucleus mutants also experience little to no perturbation as a result of mutation in the hydrophobic core. Put succinctly, mutations in the core affect the backbone and side-chain dynamics of the loop, and binding at the loop results in chemical shift perturbations in part of the core, but approximately the first 20 N-terminal residues are blind to both types of perturbations. To our knowledge, this is the first example of such an obvious pattern of correspondence between responses (or lack thereof) to structural and dynamical perturbations. It thus appears that perturbations in CI2, whether structural or dynamical, are directed toward or localized in the broader functional region, namely the reactive loop and its β-sheet support structure. CI2 and the other potato I inhibitor family members are classically non-allosteric, but the fact that distal perturbations result in effects in the functional region is reminiscent of allostery, albeit without major functional changes as the defining characteristic.

Finally, we return to the general question of what role dynamics might play in protein function. In the present work, we detected no sizeable role for dynamics in CI2 function, but that does not preclude a more dramatic connection from existing. It could simply be the case that the conservative aliphatic mutations we made in CI2's hydrophobic core were not severe enough to result in dynamical perturbations significant enough to clearly impact function, or it could also be the case that perhaps multiple mutations are necessary instead of simple single point mutations. The possible necessity of making incremental mutations to achieve a desired goal such as the dynamics-based regulation of function is reminiscent of the process

of protein evolution. In this light, the subtle functional consequences resulting from conservative mutations in CI2 could be considered to be a model of the early stages in the evolution of allostery. The foundation of functional allostery is long range communication and action at a distance – a perturbation at one site results in a functional change elsewhere in the protein. Mutations in the hydrophobic core of CI2 result in dynamical changes propagated throughout the protein, and binding to the protease target result in long range effects propagated from the reactive loop into the inhibitor's core. In summary, CI2 exhibits the capacity for long range communication, but it appears that CI2 has not evolved to harness this ability to enhance function.

From an experimental point of view, the problem with the possible need to make multiple or more severe mutations in order to elicit a more significant functional response is that this strategy reintroduces a complicating factor, namely the possibility that severe mutations will result in large scale structural deformations, in which case the consequences of altered dynamics and altered structure will no longer easily be distinguishable. This has tremendous implications for understanding the role of dynamics in larger, more complex proteins such as signal transducers and enzymes. Because the interplay of so many different factors is usually necessary to achieve overall function, it has proven quite difficult to tell whether, upon some perturbation, altered dynamics *causes* an alteration in some observable function, or whether altered dynamics passively *results* from a change in some other variable; thus, the problem of the chicken and the egg. Because of the inherent complexity of many of the most interesting and important proteins in biology and medicine, it will be difficult to tease out the general principles of functional dynamics from such systems alone. A possible worthwhile strategy for the future would therefore be to study the role of dynamics in a series of increasingly

complex or sophisticated proteins, such that minimum requirements for functional dynamics in proteins might thus become gradually apparent. This approach is explored in Chapter IV.

IV.

Initial Studies of the Role of Dynamics in the Allosteric Switching Mechanism of the Model Protein CheY

In this chapter, I turn to an initial analysis of the role of internal dynamics in the mechanism of allosteric activation of the *E. coli* chemotaxis protein Y (CheY). I first present an overview of CheY's role in chemotaxis, followed by an overview of the experimental procedures, before finally turning to the presentation and analysis of the NMR data.

IV.1. Introduction to CheY biology

Bacteria use a variety of regulatory systems to control their responses to environmental stimuli; these regulatory systems are frequently comprised of two components, a histidine kinase transmitter domain and a phosphorylatable receiver domain (*95*). Chemotaxis, the process by which bacteria such as *E. coli* alter their swimming behavior in response to chemical stimuli in the outside environment, is one such process subject to a two-component regulatory system. The chemotactic response allows bacteria to follow concentration gradients toward beneficial substances and away from potentially toxic substances. The two components of the chemotactic system are the kinase transmitter domain CheA and the phosphorylatable receiver domain CheY (*96*). In response to extracellular stimuli, CheA transfers its own phosphate group to CheY. Upon phosphorylation, CheY undergoes a subtle conformational change that increases its binding affinity for the FliM component of the bacterial flagellar motor. Increased CheY binding to FliM alters the direction of rotation of the bacterial flagella from counterclockwise to clockwise, thus altering swimming behavior

from smooth to tumbling (*97-99*). Using this system, bacteria can respond to favorable stimuli by swimming smoothly toward them and to toxic substances by tumbling to change their swimming direction, thus escaping the hazard. A schematic of the chemotactic process is shown in Figure 4.1.

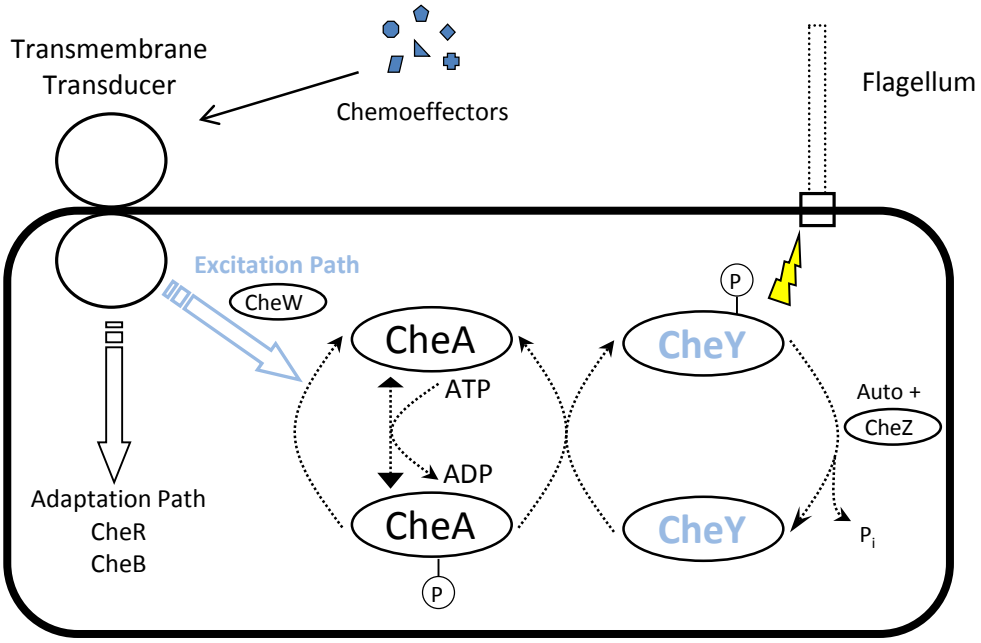


Figure 4.1. The molecular details of the chemotaxis pathway in *E. coli*. The binding of extracellular chemoeffectors to membrane transducers leads to the phosphorylation of CheA. CheA transfers its phosphate group to CheY, which then binds the flagellar motor with higher affinity, causing a switch in the sense of flagellar rotation and thus altering bacterial swimming behavior.

On a molecular level, CheY phosphorylation has subtle but long-range conformational consequences (*100, 101*). Upon phosphorylation by CheA at Asp57, the side chain of Thr87 moves closer to and forms a hydrogen bond with the phosphate group on Asp57, which allows the β4-α4 loop to collapse downward. The change in position of residues 87-89 opens a hole which the side chain of Tyr106 can fill via rotation about χ_1 ("out"→"in" transition). Tyr106 is part of the FliM binding surface comprised of the α4-β5-α5 face (*102*). In the "out" position, the Tyr106 side chain physically prevents optimal binding

contacts between CheY and FliM, but this blockage is abrogated upon rotation of Tyr106 to the "in" position, thus allowing CheY to bind to FliM with 20-fold higher affinity. The proposed method of allosteric activation involving the movements of Thr87 and Tyr106 in response to phosphorylation at Asp57 has been termed "Y-T coupling" (*100*), and the conformational transition is displayed in Figure 4.2.

Despite the physical simplicity and intuitiveness of the Y-T coupling mechanism, it does not fully explain observed CheY allostery. According to this hypothesis, the side chain of Y106 should not be able to rotate inward into the buried conformation without phosphorylation at Asp57 and concomitant movement of the β4-α4 loop, but noncanonical conformations of regions involved in Y-T coupling have been observed both experimentally (*103, 104*) and in computational simulations (*105, 106*) in the absence of the allosteric effector. This means that more must be involved in achieving CheY allosteric activation than the mechanistic conformational transition implied by the Y-T coupling hypothesis. NMR dynamics and structural techniques are in a unique position to contribute to our knowledge of allosteric activation in CheY, and in this work I have applied these techniques specifically to the allosteric transition in WT CheY.

IV.2. Materials and methods

IV.2.1. Expression and purification of WT CheY

The pet28a plasmid (Novagen) containing cDNA for WT CheY was transformed into *E. coli* BL21 (DE3) cells, and a colony containing the plasmid was used to inoculate 5 mL of LB media. The culture was grown with shaking at 37 °C and used to inoculate 45 mL of M9 media supplemented with either $^{15}NH_4Cl$ or U-^{13}C-glucose as necessary for NMR experiments. The M9 culture was allowed to shake overnight at 37 °C and was then used to

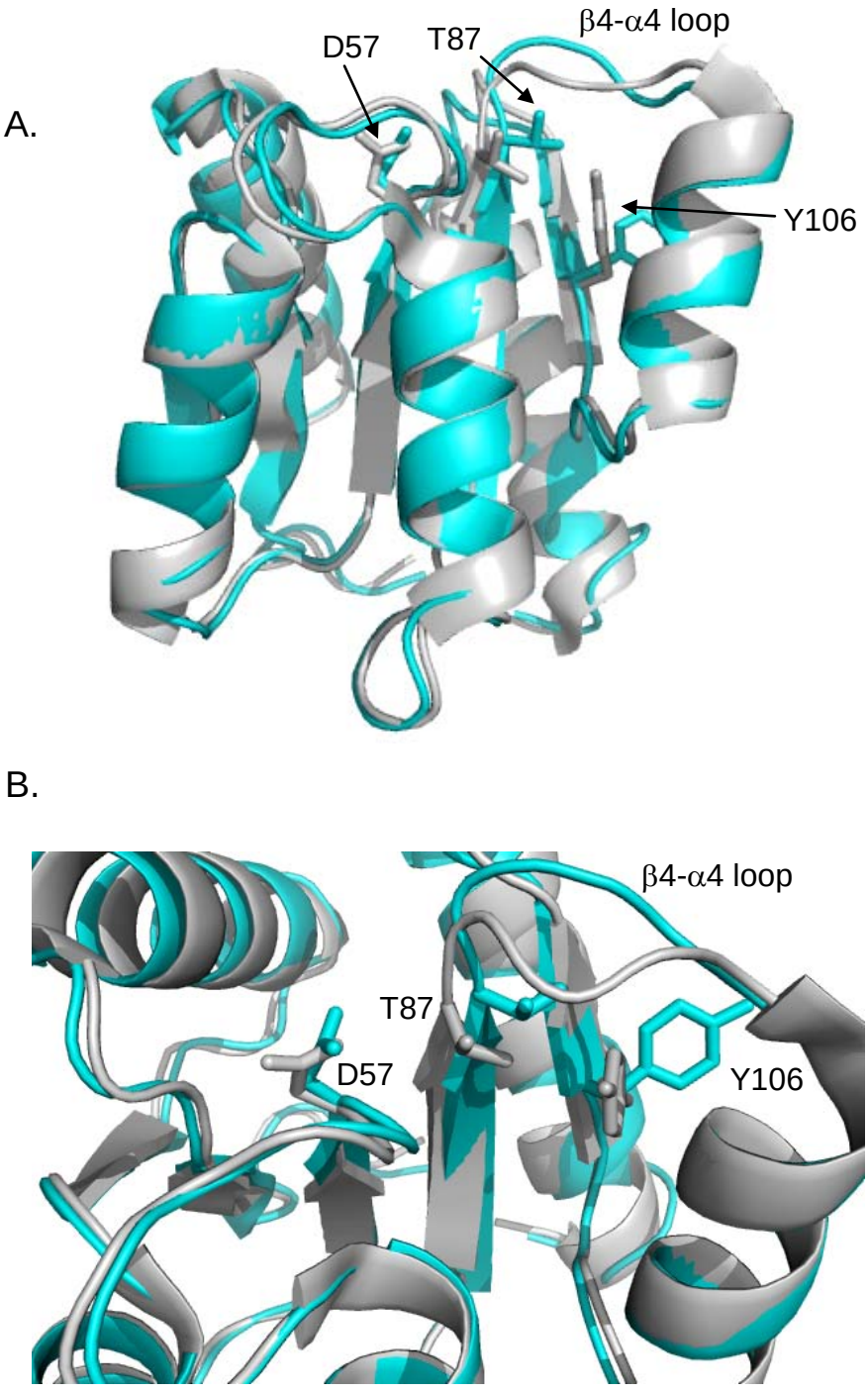


Figure 4.2. The conformational changes in response to phosphorylation at Asp57. The unphosphorylated structure (PDB 3CHY) is shown in gray, whereas the phosphorylated (activated) structure (PDB 1FQW) is shown in cyan. The important residues involved in Y-T coupling are labeled. The view in panel B is rotated 90° toward the viewer compared to the view in panel A.

inoculate 950 mL of fresh M9 media. After shaking at 37 °C until the OD_{600} reached 0.6-0.8, protein expression was induced by the addition of IPTG to a final concentration of 1 mM, and expression was allowed to proceed for 20 hours at 20 °C with shaking. The bacteria were harvested from the media by centrifugation and resuspended in Buffer A (25 mM Tris (pH 8), 10 mM $MgCl_2$). Cells were subjected to 3 rounds of freeze-thawing in a dry ice/ethanol bath, followed by 6 rounds of sonication for 5 minutes. DNA was precipitated from solution by addition of 150 μL of 5% polyethyleneimine and pelleted by centrifugation. The supernatant was retained and dialyzed overnight against 4 L of Buffer A.

Purification of CheY was achieved by passing the dialyzed supernatant over a Q Sepharose FF ion exchange column equilibrated in Buffer A. CheY was eluted from the column using a linear gradient of 0-55% Buffer A supplemented with 1.5 M NaCl. Fractions containing CheY were concentrated and passed over a G75 Superdex size exclusion column equilibrated in the NMR buffer consisting of 50 mM NaP_i (pH 7), 10 mM $MgCl_2$, and 0.02% NaN_3. Fractions containing CheY were pooled, concentrated to greater than 1 mM, and stored at 4 °C until needed.

Because CheY rapidly autodephosphorylates, the phosphorylated state is not amenable to experimental characterization in solution. Studies of the activated stated are therefore carried out on CheY activated by the phosphomimic BeF_3, which has been demonstrated to affect the structure and function of CheY and other response regulators very similarly to phosphorylation (*107*). Activation of CheY using BeF_3 was carried out as published (*100, 108*).

IV.2.2. NMR methods

Backbone and side-chain resonances from BeF_3-activated CheY were obtained by analyzing the same types of triple resonance data used to assign the resonances of CI2, namely HNCACB, CBCA(CO)NH, and (H)CCH_3-TOCSY.

Order parameters for ps-ns dynamics for backbone N-H bonds (S^2) and for the methyl symmetry axis of methyl-bearing side chains (S^2_{axis}) were calculated based on ^{15}N and ^2H spin relaxation experiments, respectively, that were performed and analyzed in the same manner reported for CI2 in the preceding chapters.

Additionally, motions on the μs-ms time scale were characterized using CPMG-based relaxation dispersion experiments for backbone amides (*74*) and side-chain methyl groups (*109*). These experiments measure R_{ex}, the contribution to R_2 from exchange processes occurring on this time scale. If an NMR-active nucleus exchanges between two different magnetic environments *A* and *B* with rate constants k_{AB} and k_{BA}, then this process will result in an observed spin-spin relaxation rate constant $R_2^{obs} = R_2^0 + \frac{p_A p_B (\Delta\omega)^2}{k_{ex}}$, where R_2^0 is the intrinsic spin-spin relaxation rate constant in the absence of exchange, $k_{ex} = k_{AB} + k_{BA}$, p_X is the fractional population of state *X*, $p_A = 1 - p_B$, and $\Delta\omega$ is the difference in chemical shift between states *A* and *B*. The preceding equation holds true when the exchange rate k_{ex} is much larger than the chemical shift difference $\Delta\omega$. I performed preliminary relaxation dispersion tests at backbone amide and methyl side-chain positions to determine whether or not these positions experience dynamics on the μs-ms time scale.

Finally, in the same manner as for CI2, side-chain ^3J couplings were measured for both the active and inactive states in order to calculate χ_1 rotamer populations and to compare rotamer populations between the two states. Significant populations of 2 or more rotameric states indicate that the side chain has significant mobility. Thus, rotameric averaging is

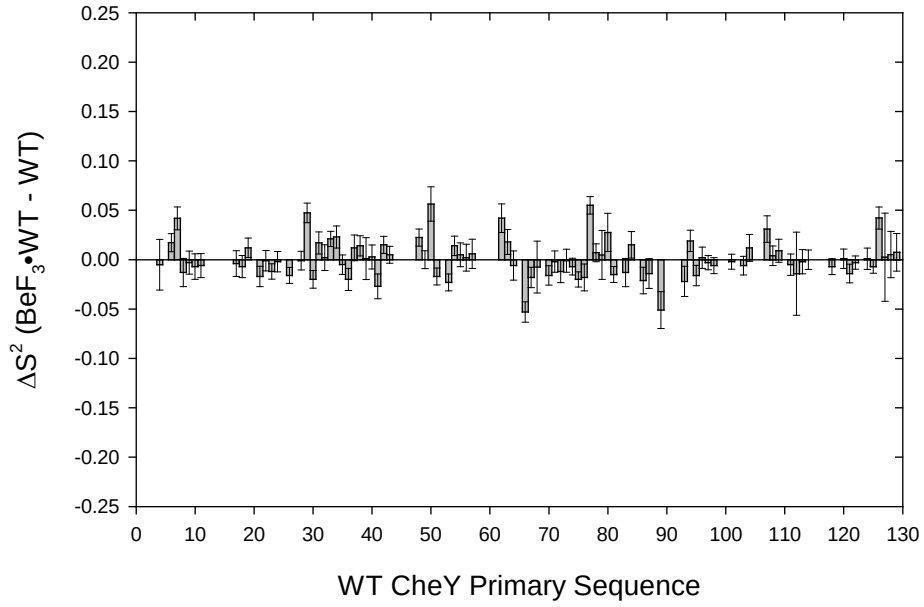
another measure of the conformational dynamics taking place at a given position in the structure, and differences in rotamer populations in the unphosphorylated and phosphorylated states indicate that phosphorylation has both a structural and dynamic effect on a given residue.

IV.3. Results and discussion

IV.3.1. Dynamics of allosteric activation on the ps-ns time scale

^{15}N relaxation experiments and calculation of backbone order parameters for unphosphorylated WT CheY were performed by Leanna McDonald, who graciously made her data available to facilitate a comparison with data I collected on BeF_3-activated WT CheY. The ps-ns backbone dynamics of allosteric activation are plotted in Figure 4.3. ΔS^2 **r**epresents the change in dynamics for a given backbone amide upon transition to the activated state. Clearly, activation has nearly no effect on backbone dynamics, as all values of ΔS^2 have a magnitude less than 0.05 and are be distributed evenly about zero.

Figure 4.3. Changes in ps-ns backbone dynamics upon allosteric activation of WT CheY calculated based on ^{15}N relaxation data.



^{2}H relaxation experiments were performed on BeF_3-activated WT CheY in order to calculate changes in methyl-bearing side-chain flexibility (ΔS^2_{axis}) in conjunction with ^{2}H relaxation data on inactive WT CheY graciously provided by Dr. Joshua Boyer. The ΔS^2_{axis} data are shown in Figure 4.4. Methyl groups with a statistically significant increase in flexibility are colored green, while an increase in rigidity is shown in cyan. The most significant change was observed for Thr87, one of the key residues involved in Y-T coupling; this side chain becomes significantly more rigid upon allosteric activation ($\Delta S^2_{axis} > 0.4$), quite possibly as a result of the formation of a hydrogen bond with the BeF_3 phosphomimic. Ala88, located at the beginning of the β4-α4 loop involved in Y-T coupling, also becomes more rigid, albeit to a lesser degree than Thr87. In panels B and C, the data from panel A are plotted onto the structure. The rigidifications (cyan) are localized in a nearly contiguous network of van der Waals interactions traversing the Y-T coupling pathway, whereas the increases in flexibility are generally localized to a distal region of the protein comprised of parts of the α1 and α5 helices. Overall, these results suggest that side-chains involved in or near the Y-T coupling trajectory rigidify upon phosphorylation. A biological role for the cluster of residues which increase in flexibility is not immediately obvious, but it is theoretically possible that an entropy compensation mechanism is at work.

IV.3.2. Dynamics of allosteric activation on the μs-ms time scale

A comprehensive analysis of μs-ms dynamics of backbone amides in inactive WT CheY has been undertaken by Leanna McDonald. Her data indicate that residues 12, 36, 38, 57, 62, 64, 86, 87, 88, 98, and 107 experience motions on this time scale (Figure 4.5A); of the

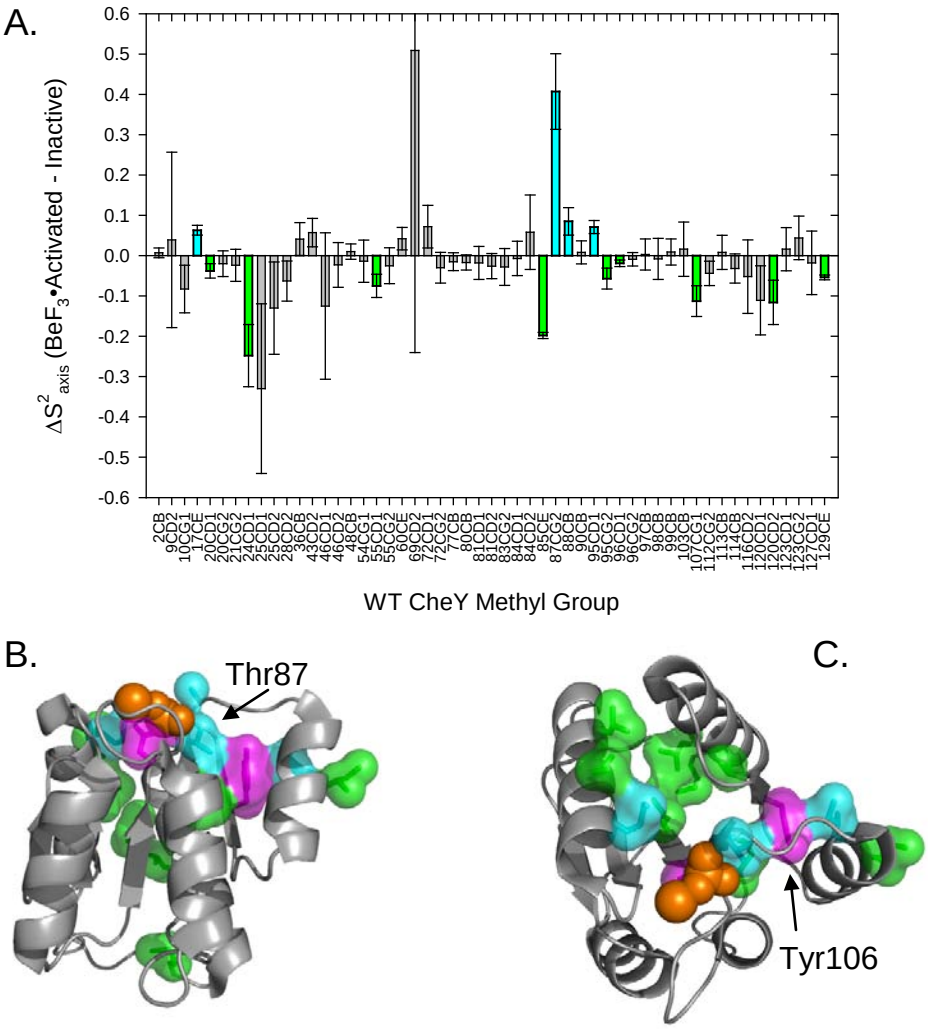


Figure 4.4. Changes in WT CheY methyl-bearing side-chain dynamics upon activation by BeF_3. Panel A shows the data for each methyl group. Panels B and C show the data from Panel A plotted onto the structure of activated WT CheY (1FQW). Panel C is rotated out of the paper toward the viewer by 90° compared to Panel B. Methyls shaded green (cyan) become more flexible (rigid) upon activation. The side chains of Asp57 and Tyr106, which cannot be analyzed using these relaxation experiments, are colored magenta, while the phosphomimic and a required magnesium ion are colored orange.

members of this list, residues 57, 86-88, and 107 are part of or near members of the Y-T coupling pathway. To determine the effect of activation by BeF_3 on the μs-ms motions at these critical positions, I performed ^{15}N relaxation dispersion test experiments on activated WT CheY. These initial results indicate the possible presence of chemical exchange at residues 38, 59, 88, 90, and 94 (Figure 4.5B). Furthermore, the test experiments suggest that these motions are greatly attenuated compared to similar locations in inactive WT CheY.

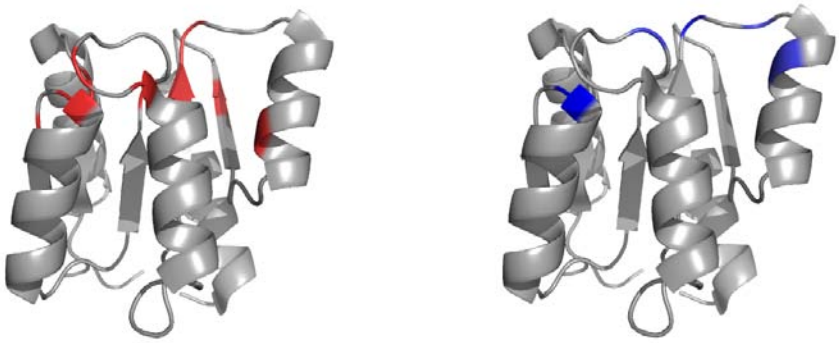


Figure 4.5. Locations of μs-ms time scale motions in inactive WT CheY (left) and BeF_3·WT CheY (right). Most of the motions indicated in the left panel have been suppressed in the right panel.

Thus, phosphorylation appears to dampen chemical exchange processes occurring in CheY; this is notable because of this dampening includes the Y-T coupling regions (surrounding T87 and Y106).

Finally, for the first time in the Lee Lab, preliminary relaxation dispersion experiments were performed to characterize μs-ms exchange processes in the methyl-bearing side chains Ala, Val, Leu, Met, and Ile in inactive and BeF_3-activated WT CheY; these experiments required the use of a specially prepared sample to insure that the methyl carbons were not directly bonded to another ^{13}C-labeled carbon (*110*). The preliminary results for inactive CheY do not indicate the presence of extensive side-chain chemical exchange. The only methyls possibly experiencing side-chain chemical exchange are Met17 and Val86. Met17 is

not surprising, because the preceding residues do not yield crosspeaks in triple resonance data, indicating the presence of some dynamical process occurring at that location, and the resonance for Met17 itself are very weak. More interesting is Val86. Because of its position next to Thr87, which cannot be analyzed using these experiments because the threonine methyl carbon does not escape being directly bonded to another ^{13}C-labeled carbon when using 1-^{13}C-glucose as the sole carbon source (*110*), the fact that Val86 appears to experience conformational exchange suggests that Thr87 might as well. The switching of the Thr87 side chain between difference environments suggests that, even in the absence of the activating phosphate group, Thr87 might not be locked into the canonical "inactive" conformation. In contrast, the same experiments performed on BeF_3-activated WT CheY did not signal the presence of side-chain conformational switching for any residues. If in fact Thr87 is mobile in the inactive state, its motions appear to be damped in the activated state, although these experiments carry no information concerning the conformation into which it may be locked.

IV.3.3. Effects of allosteric activation on side-chain structure

In a similar fashion to CI2, the measurement of $^3J_{NC'}$ and $^3J_{CC'}$ for inactive and activated WT CheY allows the calculation of χ_1 rotamer populations for residues with a γ-methyl group. The calculated populations are shown in Table 4.1. Alterations in the population of the same rotamer between the two forms of CheY indicate that activation results in some perturbation at that side chain. The side chains of most residues studied show little variation in the respective rotamer populations when the inactive and active states are compared, however the side chain of residues 10, 86, 87, and 107 show significant alterations in their respective rotamer populations upon activation. Residues 86, 87, and 107 are particularly

telling because of their involvement in or proximity to Y-T coupling. The side chains of these residues show significant rotameric averaging, meaning that all three rotameric states

	-60° Rotamer				+60° Rotamer				180° Rotamer			
	Inactive		BeF_3-Active		Inactive		BeF_3-Active		Inactive		BeF_3-Active	
Residue	*Pop*	*Error*	*Pop*	*Error*	*Pop*	*Error*	*Pop*	*Error*	*Pop*	*Error*	*Pop*	*Error*
10	0.076	0.059	0.153	0.384	0.139	0.073	0.024	0.389	0.785	0.042	0.823	0.087
20	0.878	0.027	0.848	0.060	0.123	0.032	0.152	0.092	0.000	0.041	0.000	0.110
21	0.023	0.140	0.000	0.317	0.012	0.142	-0.032	0.314	0.966	0.039	1.031	0.043
40	0.144	0.096	0.064	0.099	0.000	0.098	0.000	0.105	0.856	0.022	0.936	0.034
54	-0.062	0.112	-0.028	0.267	0.242	0.119	0.208	0.276	0.820	0.041	0.820	0.068
55	0.958	0.040	0.922	0.084	0.008	0.093	-0.009	0.314	0.034	0.101	0.087	0.325
72	0.904	0.035	0.888	0.059	0.096	0.040	0.111	0.100	0.000	0.052	0.000	0.116
83	0.000	0.151	-0.013	0.558	0.031	0.150	0.025	0.561	0.969	0.014	0.988	0.083
86	**0.101**	0.173	**0.254**	0.422	**0.064**	0.197	**0.187**	0.422	**0.836**	0.096	**0.560**	0.091
87	**0.264**	0.185	**0.401**	0.400	**0.396**	0.049	**0.273**	0.373	**0.340**	0.191	**0.325**	0.547
95	0.877	0.023	0.906	0.045	0.123	0.028	0.095	0.105	0.000	0.036	0.000	0.114
96	0.831	0.018	0.826	0.034	0.169	0.020	0.174	0.061	0.000	0.027	0.000	0.070
107	**0.409**	0.074	**0.108**	0.120	**0.613**	0.037	**0.389**	0.140	**-0.022**	0.064	**0.502**	0.072
108	0.090	0.096	0.106	0.240	0.070	0.097	0.054	0.246	0.841	0.025	0.840	0.072
112	0.247	0.090	0.257	0.145	0.752	0.017	0.743	0.052	0.000	0.092	0.000	0.154
115	0.807	0.082	0.861	0.093	0.193	0.068	0.102	0.254	0.000	0.107	0.037	0.271
123	0.908	0.039	0.927	0.065	0.092	0.049	0.072	0.155	0.000	0.063	0.000	0.168

Table 4.1. χ_1 rotamer populations for inactive and active WT CheY calculated based on ^3J scalar coupling data. Residues possibly involved in allosteric activation are in bold.

are populated to a significant degree. Rotameric averaging is a sign of structural dynamics, which indicates that even the canonical residues involved in Y-T coupling are not "locked" into a particular conformation in either the inactive or the active state

IV.4. Preliminary conclusions

Taken together, the data on ps-ns and μs-ms dynamics and the rotamer population data indicate that the mechanistic, all-or-nothing Y-T coupling hypothesis is not completely correct. While the roles of residues such as Thr87 and Tyr106 in allosteric activation are undisputable, in light of the dynamics data, these residues cannot be thought to simply and cleanly transition from their inactive to the active conformation upon phosphorylation. Rather than simply being a structural transition, allosteric activation in WT CheY is also a

dynamical perturbation having measurable effects on the ps-ns and μs-ms time scales, in addition altering the dynamics of rotameric transitions. More experiments are needed, however, to connect perturbed dynamics in the various time regimes to the functional allosteric output, increased binding affinity for FliM, and these experiments will be performed in future work.

V.

Concluding Remarks and Future Directions

The broadest theme of the work presented in this dissertation has been the relationship between protein internal dynamics, structure, and function. The work can be divided into two tiers. The first tier consisted of an NMR-based characterization of the response of structure and dynamics to conservative mutations in the hydrophobic core of chymotrypsin inhibitor 2, a protein chosen for its simplicity, in that it is neither classically allosteric nor an enzyme. The experiments revealed that CI2 experiences a robust dynamical response to subtle mutations in its hydrophobic core in the absence of significant structural perturbations. Although no clear link between CI2 dynamics and function was detected, it is still notable that perturbations have long range effects, even in this simple model system. This suggests that proteins in general might be capable of long range internal communication, even if the communication pathways are not actually harnessed for any particular functional purpose. Future studies of CI2 should involve analyzing the consequences of more severe mutations in order to determine the threshold at which the inhibitory ability starts to change significantly. Once this threshold has been determined, the dynamics of the mutants at the threshold can be analyzed to determine whether the transition from subtle to significant functional consequences is accompanied by a significant change in the dynamical response to mutation.

The second tier of the work dealt with the dynamics of allosteric activation in chemotaxis protein Y. CheY was chosen for study because, although it is an allosteric protein that

undergoes a conformational change upon binding its allosteric effector, the protein is a single domain, and the conformational transition is relatively subtle. Thus, CheY represents a single step up the chain of protein complexity from CI2. The Y-T coupling model for allosteric activation in CheY posits that CheY maintains a distinct inactive conformation until it is phosphorylated and undergoes the transition to the distinct active conformation. Analysis of the dynamics of the inactive state of CheY suggest that residues involved in Y-T coupling are mobile and flexible over a wide range of time scales, meaning that they sample a large array of conformational substates. This is inconsistent with the tenets of the Y-T coupling hypothesis, which recognizes only two discrete conformations, the transitions between which are tightly linked to the presence or absence of a phosphate group at Asp57. Analysis of activated state dynamics indicates that many of the motions present in the inactive state are damped or suppressed upon activation, suggesting that the active state samples a narrower distribution of conformational substates. Real insight into the role of dynamics in allosteric activation will be gained in future work with CheY mutants. Mutations are known which result either in constitutively active or inactive CheY without regard to the phosphorylation state. The dynamics of these mutations will be characterized and compared to WT CheY to determine whether constitutive activation or inactivation favors particular dynamical patterns.

The field of protein dynamics is still relatively young but is expanding rapidly. Clearly, a major desire of researchers in the field is to understand the physical determinants of internal dynamics and then to learn how to exploit these determinants to tune protein function. Because the field is both comparatively new and comparatively small, there is still much uncertainty as to the best way to go about studying these issues and even more uncertainty

about the best way to interpret experimental results. The strength of the present work is that it removes as many complicating layers as possible (the “messiness” of biology) in an effort to peer down as close as possible to the underlying principles. It seems only natural that those interested in the interplay between structure, dynamics, and function build up the base of knowledge concerning these relationships by examining progressively more complicated systems, and this work was designed to be serve as the first step in that chain of investigation.

References

1. **Cooper, A., and Dryden, D. T. F. (1984) Allostery without conformational change: a plausible model, *Eur. Biophys. J. 11*, 103-109.**

2. **Lee, A. L., Sharp, K. A., Kranz, J. K., Song, X. J., and Wand, A. J. (2002) Temperature dependence of the internal dynamics of a calmodulin-peptide complex, *Biochemistry 41*, 13814-13825.**

3. **Lee, A. L., and Wand, A. J. (2001) Microscopic origins of entropy, heat capacity and the glass transition in proteins, *Nature 411*, 501-504.**

4. **Mandel, A. M., Akke, M., and Palmer, A. G., 3rd. (1996) Dynamics of ribonuclease H: temperature dependence of motions on multiple time scales, *Biochemistry 35*, 16009-16023.**

5. **Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z., and Kern, D. (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair, *Nat. Struct. Mol. Biol. 11*, 945-949.**

6. **Seewald, M. J., Pichumani, K., Stowell, C., Tibbals, B. V., Regan, L., and Stone, M. J. (2000) The role of backbone conformational heat capacity in protein stability: temperature dependent dynamics of the B1 domain of Streptococcal protein G, *Protein Sci. 9*, 1177-1193.**

7. **Igumenova, T. I., Lee, A. L., and Wand, A. J. (2005) Backbone and side chain dynamics of mutant calmodulin-peptide complexes, *Biochemistry 44*, 12627-12639.**

8. **Boyer, J. A., and Lee, A. L. (2008) Monitoring aromatic picosecond to nanosecond dynamics in proteins via ^{13}C relaxation: expanding perturbation mapping of the rigidifying core mutation, V54A, in eglin c, *Biochemistry 47*, 4876-4886.**

9. **Clarkson, M. W., Gilmore, S. A., Edgell, M. H., and Lee, A. L. (2006) Dynamic coupling and allosteric behavior in a nonallosteric protein, *Biochemistry 45*, 7693-7699.**

10. **Clarkson, M. W., and Lee, A. L. (2004) Long-range dynamic effects of point mutations propagate through side chains in the serine protease inhibitor eglin c, *Biochemistry 43*, 12448-12458.**

11. **Whitley, M. J., Zhang, J., and Lee, A. L. (2008) Hydrophobic core mutations in CI2 globally perturb fast side-chain dynamics similarly without regard to position, *Biochemistry 47*, 8566-8576.**

12. Mittermaier, A., and Kay, L. E. (2004) The response of internal dynamics to hydrophobic core mutations in the SH3 domain from the Fyn tyrosine kinase, *Protein Sci. 13*, 1088-1099.

13. Fuentes, E. J., Gilmore, S. A., Mauldin, R. V., and Lee, A. L. (2006) Evaluation of Energetic and Dynamic Coupling Networks in a PDZ Domain Protein, *J. Mol. Biol. 364*, 337-351.

14. Mayer, K. L., Earley, M. R., Gupta, S., Pichumani, K., Regan, L., and Stone, M. J. (2003) Covariation of backbone motion throughout a small protein domain, *Nat. Struct. Biol. 10*, 962-965.

15. Kay, L. E., Muhandiram, D. R., Farrow, N. A., Aubin, Y., and Forman-Kay, J. D. (1996) Correlation between dynamics and high affinity binding in an SH2 domain interaction, *Biochemistry 35*, 361-368.

16. Frederick, K. K., Marlow, M. S., Valentine, K. G., and Wand, A. J. (2007) Conformational entropy in molecular recognition by proteins, *Nature 448*, 325-329.

17. Namanja, A. T., Peng, T., Zintsmaster, J. S., Elson, A. C., Shakour, M. G., and Peng, J. W. (2007) Substrate recognition reduces side-chain flexibility for conserved hydrophobic residues in human Pin1, *Structure 15*, 313-327.

18. Fuentes, E. J., Der, C. J., and Lee, A. L. (2004) Ligand-dependent dynamics and intramolecular signaling in a PDZ domain, *J. Mol. Biol. 335*, 1105-1115.

19. Boehr, D. D., McElheny, D., Dyson, H. J., and Wright, P. E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis, *Science 313*, 1638-1642.

20. McPhalen, C. A., and James, M. N. (1987) Crystal and molecular structure of the serine proteinase inhibitor CI-2 from barley seeds, *Biochemistry 26*, 261-269.

21. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding, *J. Mol. Biol. 254*, 260-288.

22. Shakhnovich, E., Abkevich, V., and Ptitsyn, O. (1996) Conserved residues and the mechanism of protein folding, *Nature 379*, 96-98.

23. Neira, J. L., Itzhaki, L. S., Otzen, D. E., Davis, B., and Fersht, A. R. (1997) Hydrogen exchange in chymotrypsin inhibitor 2 probed by mutagenesis, *J. Mol. Biol. 270*, 99-110.

24. **Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, *J. Am. Chem. Soc. 104*, 4546-4559.**

25. **Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results, *J. Am. Chem. Soc. 104*, 4559-4570.**

26. **Abragam, A. (1961) *Principles of Nuclear Magnetism*, Clarendon Press, Oxford.**

27. **Millet, O., Muhandiram, D. R., Skrynnikov, N. R., and Kay, L. E. (2002) Deuterium spin probes of side-chain dynamics in proteins. 1. Measurement of five relaxation rates per deuteron in ^{13}C-labeled and fractionally ^{2}H-enriched proteins in solution, *J. Am. Chem. Soc. 124*, 6439-6448.**

28. **Mittermaier, A., and Kay, L. E. (1999) Measurement of methyl ^{2}H quadrupolar couplings in oriented proteins. How uniform is the quadrupolar coupling constant?, *J. Am. Chem. Soc. 121*, 10608-10613.**

29. **Muhandiram, D. R., Yamazaki, T., Sykes, B. D., and Kay, L. E. (1995) Measurement of ^{2}H T_1 and $T_{1\rho}$ Relaxation Times in Uniformly ^{13}C-Labeled and Fractionally ^{2}H-Labeled Proteins in Solution, *J. Am. Chem. Soc. 117*, 11536-11544.**

30. **Zhu, L. Y., Kemple, M. D., Landy, S. B., and Buckley, P. (1995) Effect of Dipolar Cross Correlation on Model-Free Motional Parameters Obtained from ^{13}C Relaxation in AX_2 Systems, *J. Magn. Reson. Ser. B 109*, 19-30.**

31. **Werbelow, L. G., and Grant, D. M. (1975) Proton-decoupled carbon-13 relaxation in $^{13}CH_2$ and $^{13}CH_3$ spin systems, *J. Chem. Phys. 63*, 4742-4749.**

32. **Yang, D., Mittermaier, A., Mok, Y. K., and Kay, L. E. (1998) A study of protein side-chain dynamics from new ^{2}H auto-correlation and ^{13}C cross-correlation NMR experiments: application to the N-terminal SH3 domain from drk, *J. Mol. Biol. 276*, 939-954.**

33. **Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1995) Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution, *Proc. Natl. Acad. Sci. USA 92*, 9279-9283.**

34. **Bax, A., and Grishaev, A. (2005) Weak alignment NMR: a hawk-eyed view of biomolecular structure, *Curr. Opin. Struct. Biol. 15*, 563-570.**

35. **Tjandra, N., and Bax, A. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, *Science 278*, 1111-1114.**

36. Karplus, M. (1959) Contact Electron-Spin Coupling of Nuclear Magnetic Moments, *J. Chem. Phys. 30*, 11-15.

37. Hennig, M., Bermel, W., Spencer, A., Dobson, C. M., Smith, L. J., and Schwalbe, H. (1999) Side-chain conformations in an unfolded protein: χ_1 distributions in denatured hen lysozyme determined by heteronuclear ^{13}C, ^{15}N NMR spectroscopy, *J. Mol. Biol. 288*, 705-723.

38. Schnell, J. R., Dyson, H. J., and Wright, P. E. (2004) Effect of cofactor binding and loop conformation on side chain methyl dynamics in dihydrofolate reductase, *Biochemistry 43*, 374-383.

39. Grzesiek, S., Vuister, G. W., and Bax, A. (1993) A simple and sensitive experiment for measurement of J_{CC} couplings between backbone carbonyl and methyl carbons in isotopically enriched proteins, *J. Biomol. NMR 3*, 487-493.

40. Vuister, G. W., Wang, A. C., and Bax, A. (1993) Measurement of three-bond nitrogen-carbon *J*-couplings in proteins uniformly enriched in ^{15}N and ^{13}C, *J. Am. Chem. Soc. 115*, 5334-5335.

41. Bocharov, E. V., Korzhnev, D. M., Blommers, M. J., Arvinte, T., Orekhov, V. Y., Billeter, M., and Arseniev, A. S. (2002) Dynamics-modulated biological activity of transforming growth factor beta3, *J. Biol. Chem. 277*, 46273-46279.

42. Chang, C. E., McLaughlin, W. A., Baron, R., Wang, W., and McCammon, J. A. (2008) Entropic contributions and the influence of the hydrophobic environment in promiscuous protein-protein association, *Proc. Natl. Acad. Sci. USA 105*, 7456-7461.

43. D'Aquino, J. A., Gomez, J., Hilser, V. J., Lee, K. H., Amzel, L. M., and Freire, E. (1996) The magnitude of the backbone conformational entropy change in protein folding, *Proteins 25*, 143-156.

44. Johnson, E., Chazin, W. J., and Rance, M. (2006) Effects of calcium binding on the side-chain methyl dynamics of calbindin D9k: a 2H NMR relaxation study, *J. Mol. Biol. 357*, 1237-1252.

45. Stone, M. J. (2001) NMR relaxation studies of the role of conformational entropy in protein stability and ligand binding, *Acc. Chem. Res. 34*, 379-388.

46. Wand, A. J. (2001) Dynamic activation of protein function: a view emerging from NMR spectroscopy, *Nat. Struct. Biol. 8*, 926-931.

47. Wrabl, J. O., Shortle, D., and Woolf, T. B. (2000) Correlation between changes in nuclear magnetic resonance order parameters and conformational entropy:

molecular dynamics simulations of native and denatured staphylococcal nuclease, *Proteins 38*, 123-133.

48. Yang, D., and Kay, L. E. (1996) Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding, *J. Mol. Biol. 263*, 369-382.

49. Zidek, L., Novotny, M. V., and Stone, M. J. (1999) Increased protein backbone conformational entropy upon hydrophobic ligand binding, *Nat. Struct. Biol. 6*, 1118-1121.

50. Li, L., Uversky, V. N., Dunker, A. K., and Meroueh, S. O. (2007) A computational investigation of allostery in the catabolite activator protein, *J. Am. Chem. Soc. 129*, 15668-15676.

51. Popovych, N., Sun, S., Ebright, R. H., and Kalodimos, C. G. (2006) Dynamically driven protein allostery, *Nat. Struct. Mol. Biol. 13*, 831-838.

52. Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E., and Kern, D. (2005) Intrinsic dynamics of an enzyme underlies catalysis, *Nature 438*, 117-121.

53. Watt, E. D., Shimada, H., Kovrigin, E. L., and Loria, J. P. (2007) The mechanism of rate-limiting motions in enzyme function, *Proc. Natl. Acad. Sci. USA 104*, 11981-11986.

54. Ming, D., and Bruschweiler, R. (2004) Prediction of methyl-side chain dynamics in proteins, *J. Biomol. NMR 29*, 363-368.

55. Zhang, F., and Bruschweiler, R. (2002) Contact model for the prediction of NMR N-H order parameters in globular proteins, *J. Am. Chem. Soc. 124*, 12654-12655.

56. Svendsen, I., Jonassen, I., Hejgaard, J., and Boisen, S. (1980) Amino acid sequence homology between a serine protease inhibitor from barley and potato I inhibitor, *Carlsberg Res. Commun. 45*, 389-395.

57. Dokholyan, N. V., Li, L., Ding, F., and Shakhnovich, E. I. (2002) Topological determinants of protein folding, *Proc. Natl. Acad. Sci. USA 99*, 8637-8641.

58. Hilser, V. J., Dowdy, D., Oas, T. G., and Freire, E. (1998) The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble, *Proc. Natl. Acad. Sci. USA 95*, 9903-9908.

59. Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity., *J. Magn. Reson. Ser. A 103*, 203-216.

60. Uhrín, D., Uhrínová, S., Leadbeater, C., Nairn, J., Price, N. C., and Barlow, P. N. (2000) 3D $HCCH_3$-TOCSY for resonance assignment of methyl-containing side chains in ^{13}C-labeled proteins, *J. Magn. Reson. 142*, 288-293.

61. Neri, D., Szyperski, T., Otting, G., Senn, H., and Wüthrich, K. (1989) Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional 13C labeling, *Biochemistry 28*, 7510-7516.

62. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277-293.

63. Johnson, B. A., and Blevins, R. A. (1994) NMRView: a computer program for the visualization and analysis of NMR data, *J. Biomol. NMR 4*, 603-614.

64. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation, *Biochemistry 33*, 5984-6003.

65. Lee, A. L., Flynn, P. F., and Wand, A. J. (1999) Comparison of ^{2}H and ^{13}C NMR Relaxation Techniques for the Study of Protein Methyl Group Dynamics in Solution, *J. Am. Chem. Soc. 121*, 2891-2902.

66. Jarymowycz, V. A., and Stone, M. J. (2006) Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences, *Chem. Rev. 106*, 1624-1671.

67. Chen, J., Brooks, C. L., III, and Wright, P. E. (2004) Model-free analysis of protein dynamics: assessment of accuracy and model selection protocols based on molecular dynamics simulation, *J. Biomol. NMR 29*, 243-257.

68. Chou, J. J., Gaemers, S., Howder, B., Louis, J. M., and Bax, A. (2001) A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles, *J. Biomol. NMR 21*, 377-382.

69. Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra, *J. Magn. Reson. 131*, 373-378.

70. **Permi, P. (2003) Measurement of residual dipolar couplings from $^{1}H_{\alpha}$ to $^{13}C_{\alpha}$ and ^{15}N using a simple HNCA-based experiment, *J. Biomol. NMR 27,* 341-349.**

71. **Valafar, H., and Prestegard, J. H. (2004) REDCAT: A Residual Dipolar Coupling Analysis Tool, *J. Magn. Reson. 167,* 228-241.**

72. **Luginbuhl, P., Pervushin, K. V., Iwai, H., and Wuthrich, K. (1997) Anisotropic molecular rotational diffusion in 15N spin relaxation studies of protein mobility, *Biochemistry 36,* 7305-7312.**

73. **Lee, L. K., Rance, M., Chazin, W. J., and Palmer, A. G., 3rd. (1997) Rotational diffusion anisotropy of proteins from simultaneous analysis of 15N and 13C alpha nuclear spin relaxation, *Journal of biomolecular NMR 9,* 287-298.**

74. **Loria, J. P., Rance, M., and Palmer, A. G., 3rd. (1999) A Relaxation-Compensated Carr-Purcell-Meiboom-Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy, *J. Am. Chem. Soc. 121,* 2331-2332.**

75. **Best, R. B., and Vendruscolo, M. (2006) Structural interpretation of hydrogen exchange protection factors in proteins: characterization of the native state fluctuations of CI2, *Structure 14,* 97-106.**

76. **Igumenova, T. I., Frederick, K. K., and Wand, A. J. (2006) Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution, *Chem. Rev. 106,* 1672-1699.**

77. **Skrynnikov, N. R., Millet, O., and Kay, L. E. (2002) Deuterium spin probes of side-chain dynamics in proteins. 2. Spectral density mapping and identification of nanosecond time-scale side-chain motions, *J Am Chem Soc 124,* 6449-6460.**

78. **Millet, O., Mittermaier, A., Baker, D., and Kay, L. E. (2003) The effects of mutations on motions of side-chains in protein L studied by 2H NMR dynamics and scalar couplings, *J. Mol. Biol. 329,* 551-563.**

79. **Chou, J. J., Case, D. A., and Bax, A. (2003) Insights into the Mobility of Methyl-Bearing Side Chains in Proteins from $^{3}J_{CC}$ and $^{3}J_{CN}$ Couplings, *J. Am. Chem. Soc. 125,* 8959-8966.**

80. **Hu, H., Hermans, J., and Lee, A. L. (2005) Relating side-chain mobility in proteins to rotameric transitions: insights from molecular dynamics simulations and NMR, *J. Biomol. NMR 32,* 151-162.**

81. **Best, R. B., Rutherford, T. J., Freund, S. M., and Clarke, J. (2004) Hydrophobic core fluidity of homologous protein domains: relation of side-chain dynamics to core composition and packing, *Biochemistry 43,* 1145-1155.**

82. Mittermaier, A., Davidson, A. R., and Kay, L. E. (2003) Correlation between 2H NMR side-chain order parameters and sequence conservation in globular proteins, *J. Am. Chem. Soc. 125*, 9004-9005.

83. Roesler, K. R., and Rao, A. G. (1999) Conformation and stability of barley chymotrypsin inhibitor-2 (CI-2) mutants containing multiple lysine substitutions, *Protein Eng. 12*, 967-973.

84. Radisky, E. S., Kwan, G., Karen Lu, C. J., and Koshland, D. E., Jr. (2004) Binding, proteolytic, and crystallographic analyses of mutations at the protease-inhibitor interface of the subtilisin BPN'/chymotrypsin inhibitor 2 complex, *Biochemistry 43*, 13648-13656.

85. Radisky, E. S., Lu, C. J., Kwan, G., and Koshland, D. E., Jr. (2005) Role of the intramolecular hydrogen bond network in the inhibitory power of chymotrypsin inhibitor 2, *Biochemistry 44*, 6823-6830.

86. Cai, M., Gong, Y. X., Wen, L., and Krishnamoorthi, R. (2002) Correlation of binding-loop internal dynamics with stability and function in potato I inhibitor family: relative contributions of Arg(50) and Arg(52) in Cucurbita maxima trypsin inhibitor-V as studied by site-directed mutagenesis and NMR spectroscopy, *Biochemistry 41*, 9572-9579.

87. Whitley, M. J., and Lee, A. L. (2009) Frameworks for understanding long-range intra-protein communication, *Curr. Protein Pept. Sci. 10*, 116-127.

88. Hedstrom, L. (2002) Serine protease mechanism and specificity, *Chem. Rev. 102*, 4501-4524.

89. Hedstrom, L. (1996) Trypsin: a case study in the structural determinants of enzyme specificity, *Biological chemistry 377*, 465-470.

90. Hedstrom, L., Szilagyi, L., and Rutter, W. J. (1992) Converting trypsin to chymotrypsin: the role of surface loops, *Science 255*, 1249-1253.

91. Radisky, E. S., and Koshland, D. E., Jr. (2002) A clogged gutter mechanism for protease inhibitors, *Proc. Natl. Acad. Sci. USA 99*, 10316-10321.

92. Heinz, D. W., Priestle, J. P., Rahuel, J., Wilson, K. S., and Grutter, M. G. (1991) Refined crystal structures of subtilisin novo in complex with wild-type and two mutant eglins. Comparison with other serine proteinase inhibitor complexes, *J. Mol. Biol. 217*, 353-371.

93. Frigerio, F., Coda, A., Pugliese, L., Lionetti, C., Menegatti, E., Amiconi, G., Schnebli, H. P., Ascenzi, P., and Bolognesi, M. (1992) Crystal and molecular

structure of the bovine alpha-chymotrypsin-eglin c complex at 2.0 A resolution, *J. Mol. Biol. 225*, 107-123.

94. Lee, J., Natarajan, M., Nashine, V. C., Socolich, M., Vo, T., Russ, W. P., Benkovic, S. J., and Ranganathan, R. (2008) Surface sites for engineering allosteric control in proteins, *Science 322*, 438-442.

95. Parkinson, J. S., and Kofoid, E. C. (1992) Communication modules in bacterial signaling proteins, *Annu. Rev. Genet. 26*, 71-112.

96. Bourret, R. B., Hess, J. F., Borkovich, K. A., Pakula, A. A., and Simon, M. I. (1989) Protein phosphorylation in chemotaxis and two-component regulatory systems of bacteria, *J. Biol. Chem. 264*, 7085-7088.

97. Dyer, C. M., Vartanian, A. S., Zhou, H., and Dahlquist, F. W. (2009) A Molecular Mechanism of Bacterial Flagellar Motor Switching, *J. Mol. Biol.*

98. Lowry, D. F., Roth, A. F., Rupert, P. B., Dahlquist, F. W., Moy, F. J., Domaille, P. J., and Matsumura, P. (1994) Signal transduction in chemotaxis. A propagating conformation change upon phosphorylation of CheY., *J. Biol. Chem. 269*, 26358-26362.

99. Simon, M. I., Borkovich, K. A., Bourret, R. B., and Hess, J. F. (1989) Protein phosphorylation in the bacterial chemotaxis system, *Biochimie 71*, 1013-1019.

100. Cho, H. S., Lee, S. Y., Yan, D., Pan, X., Parkinson, J. S., Kustu, S., Wemmer, D. E., and Pelton, J. G. (2000) NMR structure of activated CheY, *J. Mol. Biol. 297*, 543-551.

101. Lee, S. Y., Cho, H. S., Pelton, J. G., Yan, D., Berry, E. A., and Wemmer, D. E. (2001) Crystal structure of activated CheY. Comparison with other activated receiver domains, *J. Biol. Chem. 276*, 16425-16431.

102. Lee, S. Y., Cho, H. S., Pelton, J. G., Yan, D., Henderson, R. K., King, D. S., Huang, L., Kustu, S., Berry, E. A., and Wemmer, D. E. (2001) Crystal structure of an activated response regulator bound to its target, *Nat. Struct. Biol. 8*, 52-56.

103. Dyer, C. M., and Dahlquist, F. W. (2006) Switched or not?: the structure of unphosphorylated CheY bound to the N terminus of FliM, *Journal of bacteriology 188*, 7354-7363.

104. Simonovic, M., and Volz, K. (2001) A distinct meta-active conformation in the 1.1-A resolution structure of wild-type ApoCheY, *J. Biol. Chem. 276*, 28637-28640.

105. Formaneck, M. S., Ma, L., and Cui, Q. (2006) Reconciling the "old" and "new" views of protein allostery: a molecular simulation study of chemotaxis Y protein (CheY), ***Proteins 63,*** **846-867.**

106. Knaggs, M. H., Salsbury, F. R., Jr., Edgell, M. H., and Fetrow, J. S. (2007) Insights into correlated motions and long-range interactions in CheY derived from molecular dynamics simulations, ***Biophys. J. 92,*** **2062-2079.**

107. Yan, D., Cho, H. S., Hastings, C. A., Igo, M. M., Lee, S. Y., Pelton, J. G., Stewart, V., Wemmer, D. E., and Kustu, S. (1999) Beryllofluoride mimics phosphorylation of NtrC and other bacterial response regulators, ***Proc. Natl. Acad. Sci. USA 96,*** **14789-14794.**

108. Zhao, R., Collins, E. J., Bourret, R. B., and Silversmith, R. E. (2002) Structure and catalytic mechanism of the E. coli chemotaxis phosphatase CheZ, ***Nat. Struct. Biol. 9,*** **570-575.**

109. Lundström, P., Vallurupalli, P., Religa, T. L., Dahlquist, F. W., and Kay, L. E. (2007) A single-quantum methyl 13C-relaxation dispersion experiment with improved sensitivity, ***J. Biomol. NMR 38,*** **79-88.**

110. Lundström, P., Teilum, K., Carstensen, T., Bezsonova, I., Wiesner, S., Hansen, D. F., Religa, T. L., Akke, M., and Kay, L. E. (2007) Fractional 13C enrichment of isolated carbons using [1-13C]- or [2- 13C]-glucose facilitates the accurate measurement of dynamics at backbone Calpha and side-chain methyl positions in proteins, ***J. Biomol. NMR 38,*** **199-212.**